

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 09-090963

(43)Date of publication of application : 04.04.1997

(51)Int.Cl.

G10K 15/00

H04S 1/00

H04S 7/00

(21)Application number : 07-241284

(71)Applicant : **HITACHI LTD**

(22)Date of filing : 20.09.1995

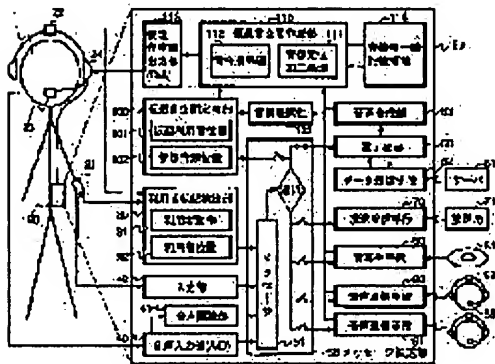
(72)Inventor : YAGAWA YUICHI
UKAI HIROMI

(54) SOUND INFORMATION PROVIDING DEVICE AND SOUND INFORMATION SELECTING METHOD

(57)Abstract:

PROBLEM TO BE SOLVED: To listen to plural sounds at the same time from different directions by locking sound image of plural pieces of sound information and putting them together.

SOLUTION: A sound image localization processing process 111 is provided in a virtual sound space generation part 110 and on the basis of a virtual user position 101 or virtual sound source position 102 set by a virtual sound space definition part 100, sound images of plural pieces of sound information are localized in a virtual sound space. Further, a user position and operation measurement part 21 and a user head direction measurement part 22 are provided and a user selects sound information intuitively and sensually. Further, a message transfer part 50 capsules the sound information that the user selects and a process for the sound information into a message 51 and transfers the message 51 to a proper processing means. Further, the processing means which is not at the virtual sound space definition part 100 generates the sound information itself. The processing means, e.g. an electronic secretary 60 is selected, a command or data from the user are sent, and the processing result is outputted as a synthesized sound by a sound synthesis part 63 to a virtual sound space generation part 110.



LEGAL STATUS

[Date of request for examination]

18.06.2002

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

BEST AVAILABLE COPY

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平9-90963

(43)公開日 平成9年(1997)4月4日

(51)Int.Cl. ⁶	識別記号	庁内整理番号	F I	技術表示箇所
G 1 0 K 15/00			G 1 0 K 15/00	M
H 0 4 S 1/00			H 0 4 S 1/00	K
				L
				F

審査請求 未請求 請求項の数18 OL (全 23 頁)

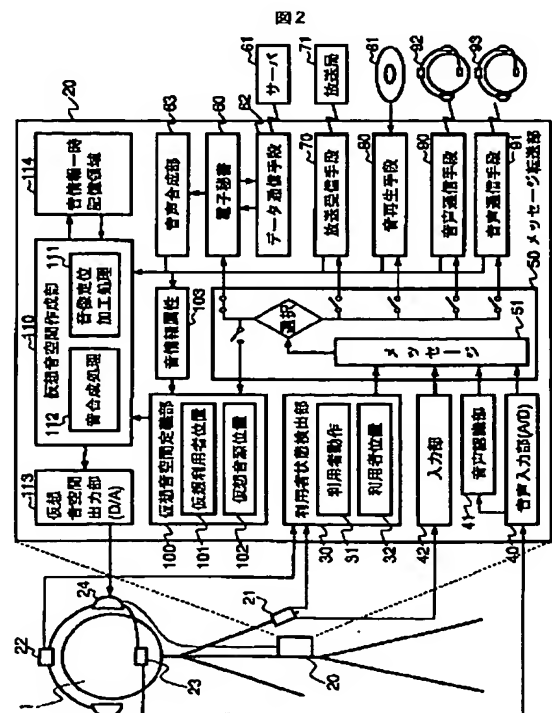
(21)出願番号	特願平7-241284	(71)出願人	000005108 株式会社日立製作所 東京都千代田区神田駿河台四丁目6番地
(22)出願日	平成7年(1995)9月20日	(72)発明者	矢川 雄一 神奈川県川崎市麻生区王禅寺1099番地株式 会社日立製作所システム開発研究所内
		(72)発明者	鶴飼 ひろみ 神奈川県川崎市麻生区王禅寺1099番地株式 会社日立製作所システム開発研究所内
		(74)代理人	弁理士 小川 勝男

(54) 【発明の名称】 音情報提供装置、及び音情報選択方法

(57) 【要約】

【目的】複数の音情報を同時に提供する手段と、利用者が欲する音情報の選択手段とを提供し、人間とコンピュータ、人間と人間との対話効率を音情報により向上させる。

【構成】利用者の位置又は動作の検出手段と、命令やデータの入力手段と、命令やデータの複数の処理手段と、処理結果を音情報として出力する手段と、利用者位置又は動作から音情報とその音情報を出力する処理の選択手段と、命令やデータを選択された処理に送る切り替えスイッチと、利用者位置又は動作に基づいて各音情報の音源を仮想音空間に配置する手段と、同じく利用者位置又は動作に基づいて仮想音空間での利用者位置を設定する手段と、仮想音空間における音源位置と利用者位置に基づいて音像を定位し、音情報が音源位置から利用者位置に向かって聞こえるかのように加工する手段と、音像定位加工後の音情報をそれぞれ出力する手段とを有する。



(2)

【特許請求の範囲】

【請求項1】複数の音情報を利用者に提供する装置であって、

前記複数の音情報を生成する処理手段と、

前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、

前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、

音像定位加工された前記複数の音情報を出力する手段と、

利用者指示を検出する利用者指示検出手段と、

前記利用者指示に基づいて、前記仮想音源位置に配置された前記音情報とその音情報を生成する処理手段を選択する音情報操作手段とを備えたことを特徴とする音情報提供装置。

【請求項2】前記利用者指示検出手段が検出する利用者指示は、利用者動作、あるいは利用者音声の少なくともいずれか一つであることを特徴とする請求項1記載の音情報提供装置。

【請求項3】前記音情報提供装置は、利用者の位置を測定する位置測定手段を備え、

前記音情報は、当該音情報が関連する事象の位置情報を属性として備え、

仮想音空間定義手段は、前記位置測定手段が測定した利用者位置と、前記音情報が表す位置に基づいて、前記利用者位置を基準とした前記音情報の仮想的な音源位置を定義することを特徴とする請求項1または2記載の音情報提供装置。

【請求項4】前記利用者指示検出手段は、検出した利用者指示から方向を表す要素を抽出し、

前記音情報操作手段は、前記方向を表す要素が示す方向の前記仮想音源位置に配置された音情報を選択することを特徴とする請求項1ないし3いずれかに記載の音情報提供装置。

【請求項5】前記音情報提供装置は、命令やデータを入力する手段を備え、

前記音情報操作手段は、前記利用者指示検出手段が検出した利用者指示または、前記入力された命令またはデータを前記選択された音情報を生成する処理手段に送ることを特徴とする請求項1ないし4いずれかに記載の音情報提供装置。

【請求項6】前記仮想音空間定義部は、前記音空間をあらかじめ定めた基準位置を中心に構成し、

前記利用者指示検出手段は、前記抽出した方向を表す要素を補正して、前記基準位置を中心とした方向に補正することを特徴とする請求項1ないし5いずれかに記載の音情報提供装置。

【請求項7】前記音情報を生成する処理手段は、音情報毎に属性を定義し、

前記仮想音空間定義手段は、あらかじめ音情報の属性毎

2

に仮想的な音空間における音源位置決定規則を定めておき、

提供される音情報の属性を前記音源位置決定規則に当てはめて音源位置を決めることを特徴とする請求項1ないし6いずれかに記載の音情報提供装置。

【請求項8】前記音情報を生成する処理手段は、指定された処理を行い、処理終了時にその処理結果、または終了通知を音情報にて出力するものであることを特徴とする請求項1ないし7いずれかに記載の音情報提供装置。

【請求項9】前記音情報は、属性として重要度または優先度を備え、

前記仮想音空間定義手段は、前記重要度または優先度が高い音情報を利用者位置の近くに定義し、前記重要度または優先度が低い音情報を利用者位置から遠くに定義するように設定することを特徴とする請求項1ないし8いずれかに記載の音情報提供装置。

【請求項10】前記音情報操作手段は、前記利用者指示に基づき、選択した音情報の音像に対し、前記音像の移動と、音量の変更と、一時的な出力の停止と、一時停止の解除と、優先出力と、優先出力からもとの状態への復帰と、音情報の出力停止のいずれかを行なうよう前記仮想音空間定義手段を制御することを特徴とする請求項1ないし9のいずれかに記載の音情報提供装置。

【請求項11】前記音出力手段の出力数は、前記音像定位加工手段が音像定位加工した音情報の数より少ない数であって、

前記音情報提供装置は、前記音像定位加工手段が音像定位加工した音情報を、前記音出力手段の出力数に合成する手段を備えたことを特徴とする請求項1ないし10のいずれかに記載の音情報提供装置。

【請求項12】複数の音情報を利用者に提供する装置であって、

前記複数の音情報を生成する処理手段と、

前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、

前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、

音像定位加工された前記複数の音情報を出力する手段と、

利用者位置を検出する利用者位置検出手段と、

前記利用者位置に基づいて、利用者位置近傍の前記仮想音源位置に配置された前記音情報とその音情報を生成する処理手段を選択する音情報操作手段とを備えたことを特徴とする音情報提供装置。

【請求項13】複数の通信相手と音声通信する装置であって、

前記複数の通信相手と通信するための通信手段と、

前記複数の通信相手と自分の仮想的な位置を音空間に定義する仮想音空間定義手段と、

50

(3)

3

前記通信手段から出力される前記複数の通信相手と自分の音声の音像を、前記定義された仮想位置に定位させる音像定位加工手段と、

音像定位加工された前記複数の音声を出力する手段と、利用者動作を検出する利用者動作検出手段と、前記利用者動作に基づいて、前記仮想位置に配置された前記通信相手を選択する音情報操作手段とを備えたことを特徴とする音声通信装置。

【請求項 1 4】複数の音情報を利用者に提供する装置であって、

前記複数の音情報を生成する処理手段と、利用者位置を検出する利用者位置検出手段と、前記利用者位置に基づいて、前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、音像定位加工された前記複数の音情報を出力する手段とを備えたことを特徴とする音情報提供装置。

【請求項 1 5】利用者を乗せ、音情報を前記利用者に提供する、移動可能な装置であって、

前記音情報を生成する処理手段と、前記移動可能な装置の位置を測定する装置位置測定手段と、前記装置の位置に基づいて、前記音空間における利用者位置を設定する仮想利用者位置設定手段と、前記音情報の仮想的な音源位置を、前記設定された利用者位置に基づいて、固定された音空間に定義する仮想音空間定義手段と、

前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、前記音像定位加工された前記音情報を出力する手段とを備えたことを特徴とする音情報提供装置。

【請求項 1 6】前記音情報提供装置は、利用者指示を検出し、検出した利用者指示から方向を表す要素を抽出する利用者指示検出手段と、前記方向を表す要素が示す方向の前記仮想音源位置に配置された音情報を選択する音情報操作手段とを備えたことを特徴とする請求項 1 4 または 1 5 に記載の音情報提供装置。

【請求項 1 7】複数の音情報の仮想的な音源位置を音空間に定義し、

前記音情報の音像を、前記定義された仮想音源位置に定位させ、前記定義された仮想音源位置から、前記複数の音情報を出力し、前記仮想音源を選択する利用者指示に基づいて、前記選択された仮想音源から出力される音情報を選択することを特徴とする音情報選択方法。

【請求項 1 8】複数の音情報の仮想的な音源位置を音空間に定義し、

4

前記音情報の音像を、前記定義された仮想音源位置に定位させ、

前記定義された仮想音源位置から、前記複数の音情報を出力し、

前記音空間を移動する利用者の位置に基づいて、前記利用者の位置近傍にある仮想音源から出力される音情報を選択することを特徴とする音情報選択方法。

【発明の詳細な説明】

【0001】

10 【産業上の利用分野】本発明は、音を用いて利用者に情報を提供する装置に係り、特に複数の音情報を同時に利用者に提供する音情報提供装置、及び利用者が複数の音情報のなかから欲する情報を選択する音情報選択方法に関する。

【0002】

【従来の技術】「いつでもどこでも」コンピュータを扱いたい、他人とコミュニケーションをとりたい、情報にアクセスしたいとのニーズから、携帯型情報通信システムや車載型情報通信システムなどいわゆるモバイルコンピュータへの期待が高まりつつある。特に、電子手帳やパーソナルデジタルアシスタント (Personal Digital Assistants、以下 PDA という) またはパーソナルインテリジェントコミュニケーター (Personal Intelligent Communicator、以下 PIC という) と呼ばれる携帯型コンピュータ及び、ページャや、携帯電話、自動車電話といった携帯型のコミュニケーションツールは、世の中に広く受け入れられつつある。

30 【0003】従来の携帯型情報通信システムは、入力手段として電子ペンやタッチパネル、出力手段として液晶ディスプレイ、コミュニケーション手段として無線データ通信装置などから構成されている。主な用途としては、スケジュール管理や住所録などいわゆるパーソナルインフォメーションマネジメント (Personal Information Management、PIM) 機能、情報やサービスの検索・利用、電子メールや電子会議等が挙げられる。また、エージェント (電子秘書) と呼ばれる電子的な秘書の概念に相当するプログラムが、情報検索やスケジュール管理など、利用者にとって面倒な仕事を代行する。例えば、利用者が入力したスケジュールを時間通りに伝達する。

【0004】これら既存のシステムにおいても、音情報を積極的に利用したものはない。従来のデスクトップ型のコンピュータと同様に、効果音など補助的に用いているのみである。以上、最近の携帯型情報通信システムについては、日経エレクトロニクス 1995 年 2 月 13 日号 (no. 628) pp 73~97、日経 BP 社に詳しい。

50 【0005】また、既存の車載型情報通信システムは、入力手段としてリモコンやジョイスティック、出力手段

(4)

5

として液晶ディスプレイ、自動車の位置を計測する自動車測位装置などから構成されている。目的地まで経路誘導する自動車ナビゲーション機能が主な用途である。また、将来には無線データ通信装置を内蔵し、道路交通情報の提供サービスに対応するシステムも開発される。音情報の利用については、音声認識によるコマンド入力機能、合成音声による情報出力機能を備えたシステムがある。

【0006】一方、コミュニケーションツールでは、小型ラジオなどの放送受信装置、携帯電話などの音声通信装置、携帯型CDプレーヤなどの音再生装置など、携帯型で音により情報を提供する装置は多い。また、例えば携帯型ラジオ付きCDプレーヤなどのように、複数のコミュニケーション手段を組み合わせ、一つの装置で複数の音情報にアクセスできるような装置も登場しつつある。

【0007】

【発明が解決しようとする課題】一般に、PDAなど既存の携帯型情報通信システムでは、情報出力手段として液晶ディスプレイを前提としているものが多い。このため、小形化は進みにくく、また仮に小形化されたとしても画面が小さく、そこから得る情報には限りがある。

【0008】省スペース及び携帯性という観点から考えると、携帯型情報通信システムでは、音情報による入力及び出力が有効である。例えば、携帯電話などは小形化が進み、スーツの内ポケットに入れて持ち運べるぐらいに小さいものまで登場しつつある。携帯型情報通信システムでも音による入出力を前提にすると、同程度まで小形化できる可能性がある。

【0009】また、車載型情報通信システムにおいても、音による情報出力が望ましい。経路情報を音声で伝えてくれると、視線を情報表示画面に移す必要もなくなる。

【0010】しかし、音情報は時系列データであり、時間軸に依存する度合が視覚情報に比べて高いため、直感性や一覧性に劣るという問題がある。このため、利用者が音情報を用いて情報通信システムと対話する場合は、逐次処理が前提となり、効率が悪い。

【0011】従って、本発明の目的は、携帯型、車載型をはじめとする各種情報処理装置において、音による情報の提供を行う装置、あるいは方法を提供し、音情報による人間とコンピュータとの対話及び人間と人間との対話の効率を向上させることである。

【0012】また、本発明の具体的な目的は、同時に提供されている複数の音情報から利用者が欲する音情報を選択する方法、装置を提供することである。

【0013】また、本発明の他の具体的な目的は、複数の音情報を提供する処理手段に対し、命令、データを入力する方法、装置を提供することである。

【0014】また、本発明の他の目的は、目が不自由な

6

人向けの情報提供装置を提供することである。

【0015】

【課題を解決するための手段】上記目的を達成するために本発明は、複数の音情報を利用者に提供する装置であって、前記複数の音情報を生成する処理手段と、前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、音像定位加工された前記複数の音情報を出力する手段と、利用者指示を検出する利用者指示検出手段と、前記利用者指示に基づいて、前記仮想音源位置に配置された前記音情報とその音情報を生成する処理手段を選択する音情報操作手段とを備えたものである。

【0016】また、前記利用者指示検出手段が検出する利用者指示は、利用者動作、あるいは利用者音声の少なくともいずれか一つであることを特徴とするものである。

【0017】また、前記音情報提供装置は、利用者の位置を測定する位置測定手段を備え、前記音情報は、当該音情報が関連する事象の位置情報を属性として備え、仮想音空間定義手段は、前記位置測定手段が測定した利用者位置と、前記音情報が表す位置に基づいて、前記利用者位置を基準とした前記音情報の仮想的な音源位置を定義することを特徴とするものである。

【0018】また、前記利用者指示検出手段は、検出した利用者指示から方向を表す要素を抽出し、前記音情報操作手段は、前記方向を表す要素が示す方向の前記仮想音源位置に配置された音情報を選択することを特徴とするものである。

【0019】また、前記音情報提供装置は、命令やデータを入力する手段を備え、前記音情報操作手段は、前記利用者指示検出手段が検出した利用者指示または、前記入力された命令またはデータを前記選択された音情報を生成する処理手段に送ることを特徴とするものである。

【0020】また、前記仮想音空間定義部は、前記音空間をあらかじめ定めた基準位置を中心に構成し、前記利用者指示検出手段は、前記抽出した方向を表す要素を補正して、前記基準位置を中心とした方向に補正することを特徴とするものである。

【0021】また、前記音情報を生成する処理手段は、音情報毎に属性を定義し、前記仮想音空間定義手段は、あらかじめ音情報の属性毎に仮想的な音空間における音源位置決定規則を定めておき、提供される音情報の属性を前記音源位置決定規則に当てはめて音源位置を決めることを特徴とするものである。

【0022】また、前記音情報を生成する処理手段は、指定された処理を行い、処理終了時にその処理結果、または終了通知を音情報にて出力するものであることを特徴とするものである。

【0023】また、前記音情報は、属性として重要度ま

50

(5)

7

たは優先度を備え、前記仮想音空間定義手段は、前記重要度または優先度が高い音情報を利用者位置の近くに定義し、前記重要度または優先度が低い音情報を利用者位置から遠くに定義するように設定するものである。

【0024】また、前記音情報操作手段は、前記利用者指示に基づき、選択した音情報の音像に対し、前記音像の移動と、音量の変更と、一時的な出力の停止と、一時停止の解除と、優先出力と、優先出力からもとの状態への復帰と、音情報の出力停止のいずれかを行なうよう前記仮想音空間定義手段を制御するものである。

【0025】また、前記音出力手段の出力数は、前記音像定位加工手段が音像定位加工した音情報の数より少ない数であって、前記音情報提供装置は、前記音像定位加工手段が音像定位加工した音情報を、前記音出力手段の出力数に合成する手段を備えたことを特徴とするものである。

【0026】また、本発明は、複数の音情報を利用者に提供する装置であって、前記複数の音情報を生成する処理手段と、前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、音像定位加工された前記複数の音情報を出力する手段と、利用者位置を検出する利用者位置検出手段と、前記利用者位置に基づいて、利用者位置近傍の前記仮想音源位置に配置された前記音情報とその音情報を生成する処理手段を選択する音情報操作手段とを備えたことを特徴とするものである。

【0027】また、本発明は、複数の通信相手と音声通信する装置であって、前記複数の通信相手と通信するための通信手段と、前記複数の通信相手と自分の仮想的な位置を音空間に定義する仮想音空間定義手段と、前記通信手段から出力される前記複数の通信相手と自分の音声の音像を、前記定義された仮想位置に定位させる音像定位加工手段と、音像定位加工された前記複数の音声出力する手段と、利用者動作を検出する利用者動作検出手段と、前記利用者動作に基づいて、前記仮想位置に配置された前記通信相手を選択する音情報操作手段とを備えたことを特徴とするものである。

【0028】また、本発明は、複数の音情報を利用者に提供する装置であって、前記複数の音情報を生成する処理手段と、利用者位置を検出する利用者位置検出手段と、前記利用者位置に基づいて、前記複数の音情報の仮想的な音源位置を音空間に定義する仮想音空間定義手段と、前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、音像定位加工された前記複数の音情報を出力する手段とを備えたことを特徴とするものである。

【0029】また、本発明は、利用者を乗せ、音情報を前記利用者に提供する、移動可能な装置であって、前記音情報を生成する処理手段と、前記移動可能な装置の位

8

置を測定する装置位置測定手段と、前記装置の位置に基づいて、前記音空間における利用者位置を設定する仮想利用者位置設定手段と、前記音情報の仮想的な音源位置を、前記設定された利用者位置に基づいて、固定された音空間に定義する仮想音空間定義手段と、前記音情報の音像を、前記定義された仮想音源位置に定位させる音像定位加工手段と、前記音像定位加工された前記音情報を出力する手段とを備えたことを特徴とするものである。

【0030】また、前記音情報提供装置は、利用者指示を検出し、検出した利用者指示から方向を表す要素を抽出する利用者指示検出手段と、前記方向を表す要素が示す方向の前記仮想音源位置に配置された音情報を選択する音情報操作手段とを備えたことを特徴とするものである。

【0031】また、本発明は、複数の音情報の仮想的な音源位置を音空間に定義し、前記音情報の音像を、前記定義された仮想音源位置に定位させ、前記定義された仮想音源位置から、前記複数の音情報を出力し、前記仮想音源を選択する利用者指示に基づいて、前記選択された仮想音源から出力される音情報を選択することを特徴とするものである。

【0032】また、本発明は、複数の音情報の仮想的な音源位置を音空間に定義し、前記音情報の音像を、前記定義された仮想音源位置に定位させ、前記定義された仮想音源位置から、前記複数の音情報を出力し、前記音空間を移動する利用者の位置に基づいて、前記利用者の位置近傍にある仮想音源から出力される音情報を選択することを特徴とするものである。

【0033】
【作用】一般に、コンピュータと人間との対話の効率を向上させる一つの方法としては、複数の仕事を同時に行う並行処理があげられる。既存のコンピュータシステムでは、並行処理を行うための手段として、「マルチウィンドウ」が用意されている。また、利用者が「マルチウィンドウ」ごとに分けられた処理を選択する手段として、マウスなどの「ポインティングデバイス」が用意されている。利用者は「マルチウィンドウ」ごとに処理を割り当てて、必要に応じてポインティングすることにより処理を選択する。利用者は意識するしないにかかわらず、平行的に仕事を進めることができ、この結果、人間とコンピュータとの対話の効率も向上する。音の世界にも同様の仕組みを導入すれば、対話の効率も上がり、さらには、電話や放送など既存のコミュニケーション手段をそれぞれ「マルチウィンドウ」として組み合わせたシステムも可能になる。

【0034】実世界において、無音という状況はありえない。人間は常に何らかの音を無意識のうちに聞き分けており、自分に必要な音情報だけを選択して聞いている。騒がしいカクテルパーティであっても自分の興味ある話題が会話にのぼると、たとえ離れたところにも

(6)

9

自然に耳に入る。音響心理学において、先の人間の特性はこの現象にちなんでカクテルパーティ効果と呼ばれている。

【0035】コンピュータが作り出す仮想的な音の世界で絶えず多様な音が聞こえていても、カクテルパーティ効果により利用者は音情報を取捨選択できる。そこで、まず音源を仮想的な音の空間に配置することによって音情報を区別しやすくする。複数の音情報が同時に聞こえたとしても、聞こえてくる方向がそれぞれ違うため、利用者は音情報を識別しやすい。この仮想的な音空間に配置した音源が「ウィンドウ」に相当する。

【0036】更に、複数の音情報の中から利用者が希望のものを選択する手段を備える。基本的に、空間内で方向を示せるものであれば良く、本発明では利用者の位置または動作により、音情報を選択する。例えば、利用者がある地点に移動すると、その場所に対応付けられた音情報が聞こえる。あるいは、音が聞こえてくる方向を指し示すことにより、音情報を選択する。この利用者の位置または動作により音情報を選択する手段が「マウス」に相当する。

【0037】本発明では、仮想的な空間に情報を処理する手段を割り当てて、処理結果は音で返す。音が聞こえてくる方向によって処理内容を判断できるので、利用者は処理結果に対してすばやく応答できる。また、電話や放送など従来のコミュニケーション手段も同じ音空間に割り付ける。利用者は仮想的な音空間から必要な処理を次々と選択することによって平行して仕事を進めることができ、結果的に人間とコンピュータとの対話及び人間と人間との対話の効率も向上する。

【0038】

【実施例】以下、図面を参照して本発明の一実施例を説明する。

【0039】＜発明の概要説明＞図1は、本発明の特徴を概念的に表す図である。まず、音源として、利用者1からの命令をサーバコンピュータまたはローカルなコンピュータで代行して処理してその処理結果を合成音声で出力する電子秘書2と、通信手段を用いて複数の相手と通信する電子会議3及び4と、ラジオなどの放送またはCDプレーヤなどの音再生5がある。それぞれの出力音は、あたかも仮想音空間10内のそれぞれ音像6、7、8、9に実際の音源があるかのように聞こえる。

【0040】音像とは音響的に作られた感覚上の音源のことを言い、本実施例では仮想音源と同義に用いる。利用者1は頭部や手等、身体の一部を用いてそれぞれの音像すなわち音情報を選択し、音情報に対する各種処理を実行する。利用者が行う処理には、例えば音像位置の移動、音量変更、音量最小化、音量最小化解除、一時停止、一時停止解除、出力優先（他の音源からの出力をすべて一時停止）、出力優先解除、停止などがある。音情報の内容、種類に応じて音像を配置できるので、利用者

10

は出力音が聞こえてくる方向からその内容を容易に把握できる。また、複数の音情報を同時に聞く場合だけでなく、必要に応じて他の音源からの出力量を抑えて、一つの音情報を聞くなどでもできる。

【0041】音像6～9にはそれぞれ処理手段2～5が対応付けられている。つまり、利用者1が仮想音空間10内の音像を選択することはすなわち処理手段をも選択したことになる。よって、利用者1はコマンドや音声などのデータを各処理手段に入力することができ、また処理手段を切り替えることによって平行的に仕事を進めることが可能になる。

【0042】＜装置構成概略説明＞図2は、本発明の一実施例に係る音情報提供装置の利用イメージと基本構成を表す。本装置は携帯型であり、その外部構成は本体部20と、利用者1の現在位置及び手の動作を測定する利用者位置・動作測定部21と、利用者1の頭の向きを測定する利用者頭部向き測定部22と、利用者1の音声を入力するマイク23と、本体部20が生成した仮想音空間の出力音を利用者1が聞くためのヘッドホン24とから成る。ここで、利用者位置・動作測定部21は腕時計型で、利用者1がコマンドやデータの一部を入力する手段とともに兼ね備えている。また、マイク23をヘッドホン24に接続せず、利用者位置・動作測定部21に内蔵し、必要に応じてマイクを口に近づけて音声入力する構成も考えられる。

【0043】本体部20の内部構成は、利用者位置・動作測定部21及び利用者頭部向き測定部22で測定されたデータを解析して利用者動作31及び利用者位置32を検出する利用者状態検出部30と、利用者の入力音声をA/D（アナログ・デジタル）変換する音声入力部40と、入力音声を認識してコマンドやデータの文字列に変換する音声認識部41と、本体部20または利用者位置・動作測定部21に設けられたボタン等からコマンド及びデータを入力する入力部42と、利用者が選択した音情報を判断し、その音情報の識別番号と当該音情報を出力する仮想音源の識別番号とコマンドまたはデータをメッセージ51としてカプセル化し、そのメッセージ51を適切な処理手段に転送するメッセージ転送部50と、利用者のコマンドを処理する電子秘書60と、利用者のコマンドをサーバコンピュータ61で処理するためのデータ通信手段62と、それぞれの処理結果を合成音声で出力する音声合成部63と、ラジオ等の放送局71からの電波を受信する放送受信手段70と、CD81等を再生する音再生手段80と、他の利用者92～93とコミュニケーションするための音声通信手段90～91と、前記メッセージ51の内容や各音情報の属性103に従って仮想的な音空間内における利用者位置101または音源位置102が定義されるよう規則を決める仮想音空間定義部100と、前記定義した位置に基づいて音情報を仮想的な音空間に音像定位し加工する仮想音空間

(7)

11

作成部110と、前記作成結果または個別の音情報を一時的に記憶する音情報一時記憶領域114と、音情報をD/A(デジタル・アナログ)変換してヘッドホン24から出力する仮想音空間出力部113とから成る。

【0044】本実施例では、仮想音空間作成部110に音像定位加工処理111を設けたことにより、仮想音空間定義部100で設定した仮想利用者位置101または仮想音源位置102に基づいて、複数の音情報を仮想的な音空間に音像定位できる点に特徴がある。ここで音像定位とは、音を音響的に加工し、感覚上の音源である音像を作り出すことである。この音像定位により、利用者は本装置が作り出した音像から本当に音情報が聞こえてくるように感じる。また、仮想音空間定義部100で音像の位置を調整することにより、複数の音情報がそれぞれ違った方向から聞こえてくるといった音響効果を作り出すことができる。利用者は音情報をその聞こえてくる方向から明確に区別できるので、複数の音情報の中から一つを選択して処理することが可能になる。

【0045】次に、本実施例では、利用者位置・動作測定部21や利用者頭部向き測定部22を設けたことにより、利用者が音情報を直感的にそして感覚的に選択できる点に特徴がある。具体的には、利用者状態検出部30で利用者が向いた方向や指し示した方向を検出し、その方向から聞こえてくる音情報を選択したものと判断する。また、ボタンや音声で音情報を選択することもできる。この場合、音情報が聞こえてくる方向を音声で指定したり、または方向が記載されたボタン等を使って選択する。あるいは、それぞれの音情報に名前等の識別子を付けておき、その識別子を音声やボタン等で選択する方法も考えられる。なお、これら利用者の向き、指し示した方向、発声、ボタン操作などを総称して利用者指示ということにする。

【0046】さらに、本実施例では、メッセージ転送部50で利用者が選択した音情報とそれに対する処理をメッセージ51としてカプセル化し、適切な処理手段にそのメッセージを転送する点に特徴がある。ここでいう処理手段には、仮想音空間定義部100、電子秘書60、放送受信手段70、音再生手段80、音声通信手段90及び91がある。仮想音空間定義部100は音情報が音像定位されている仮想的な音源に対する処理を実行する。例えば、先に検出された利用者位置32または身振り手振りなどの利用者動作33から仮想的な音空間における利用者位置または音源位置を設定する。すなわち、利用者は同時に聞こえる複数の音情報の中から優先して聞きたい音情報を選択したり、聞きたくない音情報を消したり、音情報の聞こえる方向やそれぞれの音量を自由に変更したりできる。

【0047】また、仮想音空間定義部100以外の処理手段では、音情報そのものを生成する。利用者が音情報を選択したということは、その音情報を生成する処理手

12

段をも選択したことになり、当該処理手段にコマンドまたはデータを送ることになる。図2の例では、電子秘書60が選択され、利用者からのコマンドまたはデータが送られ、その処理結果が音声合成部63で合成音として仮想音空間作成部110に出力されている。

【0048】<装置外觀図の説明>図5、図6は、本体部20と利用者位置・動作測定部21のそれぞれ装置外觀の一例を示す図である。画面201にはタッチパネルが重ねられていて、入力用ボタンの数を減らすためにプログラムによって画面201上に表示されたソフトウェアボタンをペン202、または、指で選択するという操作を行う。また、仮想音空間における音像の配置を画面に表示し、利用者が音像の位置をビジュアルに確認したり、表示された音像を画面上で選択するという操作も行う。

【0049】ボタン203は、モード切替や、頻繁に使うコマンド入力用のソフトウェアボタンまたは機械的なボタンである。例えば、電子会議用のコマンドを入力したいときは電子会議モードのボタンを選択し、電子秘書にコマンドを送りたいときは電子秘書モードのボタンを選択する。つまり、モードは音情報を出力する処理の種類ごとに用意されている。頻繁に使うコマンドとしては、例えば音量調節等がある。カードスロット204～206には、赤外線通信機、ラジオ受信機やパーソナル・ハンディフォン・システム(以下、PHSという)通信機等のカードを差す。利用者が必要に応じて通信手段を選択できるフレキシブルな構成となっている。

【0050】カードとして、最近普及しつつあるパーソナル・コンピュータ・メモリ・カード・インターナショナル・アソシエーション(Personal Computer Memory Card International Association、以下、PCMCIAという)対応のカードを用いる。また、図5ではスロットの数が3個となっているが、特にこの数に限定しているわけではない。

【0051】本体部20には、赤外線通信機207と有線データ通信のためのモデム208が内蔵されている。利用者は本体部20を持ち運んで利用するため、カバー209で画面201を保護する。本実施例では、画面201を主に入力的手段として用いているが、もちろん処理結果を音情報として出力するのにあわせて、画面201に処理結果をビジュアルに表示してもよい。

【0052】利用者位置・動作測定部21にも画面211が備えられている。通常は、時刻を表示しているが、利用者からの要求に応じて仮想音空間における音像の配置等を表示する。また、ボタン212は、本体部20で行えるコマンドのショートカットキーまたは数字等の簡単なデータ入力もできるようになっているので、利用者は本体部20をいちいち取り出して操作する必要がない。マイク213はヘッドホン24に具備してあるマイ

(8)

13

ク23を利用者が屋外で用いるのに不都合がある場合に用いる。利用者位置・動作測定部21は腕時計型の形状なので、バンド214で利用者の腕に固定して用いる。なお、上記画面201、211は液晶パネルを使用するのが一般的である。

【0053】＜ハードウェア構成の説明＞図2の各構成要素を実現するハードウェア構成について、図3を用いて説明する。まず、利用者位置・動作測定部21は、利用者の現在位置を測定するグローバル・ポジショニング・システム（以下、GPSという）120と、利用者の動作を測定するジャイロ121から成り、本体部20へは赤外線通信機を用いて無線でデータ送信する。GPSは周回衛星を利用した測位システムであり、最近自動車用ナビゲーションシステムの自動車測位装置として広く用いられている。各周回衛星からは衛星の経度、緯度、高さ、時刻のデータが送信され、これをGPSが受信する。そして、到達するまでにかかった時間から各衛星までの距離を計算し、GPSの現在位置を得る。測位精度を上げる工夫もなされており、現状では位置を数十メートルの誤差内で特定できる。また、小形化も進み、腕時計と同程度の大きさのGPS装置も登場するとの予測もある。GPSについては、日経エレクトロニクス（日経BP社）1995年2月13日号（no.628）pp187～194などに詳しい。

【0054】ジャイロ121は、角速度センサの一種で、内蔵の振動子の振動から角速度または回転角を検出する。利用者の腕の向きの水平方向と垂直方向の回転角をそれぞれ検出できるよう、二軸ジャイロ（または水平方向用と垂直方向用の二つのジャイロ）を用いる。当該ジャイロでは一定時間ごとに回転角を測定し、本体部20に送る。本体部20では、当該回転角の変化量から利用者の手の動きを計算する。実装については、圧電振動ジャイロ等を用いることで携帯可能な利用者位置・動作測定部21を開発できる。

【0055】利用者頭部向き測定部22もジャイロ122を用いて利用者の頭部の向いている方向を測定する。ここでも、二軸ジャイロ（または水平方向用と垂直方向用の二つのジャイロ）を用い、水平方向と垂直方向の頭部の向きを測定する。また、図2に示すように、本体部20へは有線でデータを送信している。スピーカ151は、利用者がヘッドホン24の代わりに用いる。スピーカ151については後述する。

【0056】次に、本体部20は、図4に記載した各種プログラムを実行する中央処理装置（以下、CPUという）130と、プログラムやデータを一時保存するメモリ131と、プログラムや処理パラメータを記憶しておく外部記憶装置（以下、ディスクという）132と、利用者がモード切替、コマンド及びデータを入力するボタン類141（図5、図6のボタン203、ボタン212に相当する）と、マイクから入力された音声を変換する

14

・デジタル変換するA/D変換器140と、出力音を変換するデジタル・アナログ変換するD/A変換器150と、サーバコンピュータとデータ通信するための赤外線通信機160と、ラジオ受信機161と、CDプレーヤ162と、データ通信または音声通信のためのPHS通信機163及び164と、各音情報をアナログ・デジタル変換するA/D変換器165～168とから成る。なお、図示していないが、本体部20と、利用者位置・動作測定部21には上述の液晶パネルがあり、CPU130からの指示に従って、表示を行う。

【0057】ボタン類141の例としては、電源のON/OFFスイッチ、音量調整つまみ、放送局選択用の周波数調整つまみ、CDプレーヤ制御用ボタン、アルファベットやひらがな等を入力するキャラクタボタン、電話番号等を入力する数値ボタン等がある。装置を携帯できるサイズにおさめるため、本体部20や利用者位置・動作測定部21に液晶ディスプレイとタッチパネルを備え、プログラムによって画面に表示されたグラフィカル・ユーザ・インタフェース（以下、GUIという）のボタンやメニュー（以下、ソフトウェアボタンまたは単にボタンと呼ぶ）を利用者がペン、指等で選択する構成である。画面201、211、ボタン203、ボタン212がボタン類141に相当する。タッチパネルは、感圧式、電磁誘導式、超音波式、静電誘導式などの方式が使用可能である。

【0058】また、赤外線通信機160、ラジオ受信機161、PHS通信機163、164は、カードスロット204～206に必要に応じてカードを入れ替えるといった構成を採用する。以上述べてきたハードウェア構成要素は、システムバス169で互いに連結されており、データの受け渡しや制御をこのシステムバスを通じて行う。

【0059】＜ソフトウェア構成の説明＞図4は、プログラムとデータ、及びプログラム間の制御とデータの流れを示すソフトウェア構成図である。また、図中の矩形のブロックはプログラムを表し、角が丸いブロックはデータを表す。この図から、図2で記載した利用者状態検出部30、メッセージ転送部50、仮想音空間定義部100、仮想音空間作成部110は複数のプログラムとデータから成るソフトウェアモジュールで実現されることがわかる。また、本実施例では、本体部20を携帯サイズに収めるため、音声認識部41、音声合成部63もソフトウェアで実現する。

【0060】なお、これらソフトウェアは、通常ディスク132に格納されており、必要に応じてメモリ131に読み出され、CPU130がそれを実行するものである。もちろん、メモリ131上にこれらソフトウェアがあらかじめ格納されていてもよい。

【0061】以下、プログラム制御の流れに従って各構成要素を順次説明する。入力音声データ43は、利用者

(9)

15

がマイク23を用いて入力した音声をアナログ・デジタル変換したものである。音声認識部41では、この入力音声データ43を認識し、コマンド44またはデータ45に変換する。ここで、ディスプレイやキーボード等を用いず音声情報だけで認識誤りを修正することが困難な場合は、認識率を向上させるため、利用者は単語毎に区切って発話するものとする。また、特定話者の単語音声認識に限定し、認識方式も実用化が進んでいるDPマッチングアルゴリズム等を用いる。

【0062】コマンド46とデータ47は、本体部20または利用者位置・動作測定部21に設けたボタンからの入力に相当する。先のコマンド44及び46とデータ45及び47は、それぞれメッセージ転送部50に送られる。入力音声データ43は、他の利用者と音声通信する場合にメッセージ転送部50に送られる。利用者状態検出部30には、利用者動作検出処理33と、利用者位置検出処理34がある。ある一定時間ごとに利用者動作と利用者位置を検出し、それぞれ、利用者動作管理テーブル35と利用者位置管理テーブル36に記録する。詳細は後述する。

【0063】メッセージ転送部50では、メッセージ51を作成し、適切な処理手段へ転送する。以下、メッセージ51のデータ構造及び処理方法について順に説明する。本発明では、並行処理を前提としており、並行して走るプログラムをプロセスと呼ぶ。電子秘書プログラム（以下、エージェントという）群64やハードウェア制御プログラム72などのように、利用者が陽に指定して起動するプログラムをユーザプロセスと呼び、仮想音空間の構成を変えるプログラム等、システムが起動・制御するプログラムをシステムプロセスと呼ぶ。メッセージ51は利用者が入力したコマンド及びデータを、それらを渡すべきプロセスと仮想音源の識別子とともにカプセル化したものである。そして、そのデータ構造は、メッセージ自身の識別子であるメッセージIDと、仮想音源の識別子である仮想音源ID56と、プロセスの識別子であるプロセスID57と、コマンド58と、データ59とからなる。

【0064】メッセージ51が作成されると、メッセージ転送処理55では、次にメッセージ内のプロセスIDに相当するプログラムすなわち、仮想音空間定義部100と、エージェント群64と、ハードウェア制御プログラム群72のいずれかに当該メッセージを転送する。これらのプログラムは、メモリ上に置かれたイベント駆動型のプログラムである。つまり、何らかの要求がプログラムに送られると、各プログラムはCPU130をタイムシェアして使いながら、処理を実行する。

【0065】なお、エージェント群64としては図10に示す種類のものが考えられるが、新たに利用者または第三者が作成したエージェントをシステムに導入できる。この際、エージェントの名称（プロセス名称）やコ

16

マンドジェスチャ、コマンドボイス、コマンドボタンを新たに定義しておく。エージェント群64の処理結果は、文字列として音声合成部63に送られて音声データ103として出力する。

【0066】ハードウェア制御プログラム群72としては、PHS通信機163、164を制御して電子会議を行うための電子会議プログラム、ラジオ受信機161、CDプレーヤ162、赤外線通信機160を制御するプログラムを用意しておく。処理結果、受信内容、再生内容は音情報115として出力される。

【0067】仮想音空間定義部100では、仮想的な音空間における利用者の位置と音源の位置を設定する。この設定では、利用者の位置は固定したままで仮想音源の位置を変える場合（仮想音源位置設定処理104）と、あらかじめ仮想音源が設置されている仮想音空間を利用者が自由に動きまわる場合（仮想利用者位置設定処理105）の二種類の利用場面について考える。

【0068】仮想音源位置設定処理104では、図12に示す仮想音空間管理テーブル106を用いて仮想音源の位置を定義・管理する。当該テーブルは、仮想音源IDごとに、プロセスID、出力状態、仮想音源位置からなる。プロセスIDは、それぞれの仮想音源に割り当ててあるプロセスを識別するために用いる。特に、利用者が仮想音源を選択してコマンドまたはデータを入力する際に、システムがどのプロセスに対して入力が行われたかを識別するために用いる。

【0069】また、音情報は視覚情報に比べ、継続性がない。そのため、聞き逃した音を再度聞き直すことができるように、音情報一時記憶領域114をメモリ131、またはディスク132内に用意しておく。

【0070】＜詳細な説明＞次に、上述の各構成要素の更に詳しい構成を説明する。まず、利用者動作検出処理33の動作について図7、図8を用いて説明する。図7はメモリ131に設ける利用者動作管理テーブル（以下、動作管理テーブルという）35の構成を示す図である。この動作管理テーブル35には、検出時刻とその時刻における手の向きと頭の向きとを記録する。ここでは、水平方向と垂直方向の基準面からの回転角の組で向きを表している。

【0071】図8に頭の向きを検出する場合の例を示す。垂直方向の基準面220は常に利用者1の正面にあり、回転角は利用者1が正面から向かって右方向を向いた場合に正の値をとり、左方向を向いた場合に負の値をとる。図8(a)では、利用者1の頭部の向きは、水平方向で-45度となっている。また、水平方向の基準面221は常に利用者1の耳の位置にあり、回転角は利用者1が上方向を向いた場合に正の値をとり、下方向を向いた場合に負の値をとる。図8(b)の例では、利用者1の頭部の向きは、垂直方向で+20度となっている。それぞれの回転中心については後述する。

(10)

17

【0072】図7は、図8(a)での利用者頭部の向きを記録している。すなわち、時刻が18時30分40秒から42秒までの間に、頭部の向きが水平方向に-45度だけ変化している。利用者の動作は短時間に变化する可能性があるため、1秒といった短い単位時間ごとに検出する。また、テーブル35は、利用者動作の変化を10分程度連続して記録できるように容量を決め、メモリ131に設ける。手の動きについても同様であるが、詳細は後述する。

【0073】利用者動作検出処理33は、まず、ジャイロ122で測定した頭部の水平方向と垂直方向の回転角を、バス169を介して得る。そして、ジャイロ122がとりつけてある位置と基準面との誤差を考慮し、回転角を補正して動作管理テーブル35に記録する。

【0074】図2に示すように、ジャイロ122を内蔵する利用者頭部向き測定部22が利用者1の頭部頂点に配置してある場合は、水平方向の誤差はないが、垂直方向では利用者の耳の位置とジャイロ122の位置が異なるため、耳から頭部頂上までの高さ分の誤差補正を行う。各利用者間で耳から頭部頂上までの高さの個人差は無視できるほど小さいと考えると、誤差補正処理があらかじめジャイロに組み込まれているといった構成でもよい。

【0075】手の向きを検出する場合も上記とほとんど同じ処理になる。利用者動作検出処理33では、ジャイロ121の位置と基準面との誤差を考慮し、ジャイロ121で測定した回転角を補正して動作管理テーブル35に記録する。垂直方向の基準面220は利用者の正面にあり、回転角は利用者が正面から向かって右方向に手を向けた場合に正の値をとり、左方向に向いた場合に負の値をとる。同じく、水平方向の基準面221は常に利用者の耳の位置にあり、回転角は利用者が手を上方向に向けた場合に正の値をとり、下方向に向けた場合に負の値をとる。手の動きにあわせてジャイロ121の位置が変わることによって生じる誤差については、利用者動作・位置測定部21にジャイロの位置を測定する装置を内蔵するか、または誤差を許容または予測できる範囲での手の移動しか認めないといった制限を設け、補正を可能にする。

【0076】以上の説明では、音像を三次元空間に定位させることを前提としていたが、装置を簡単かつ安価にするため、音像を水平方向にのみ定位させる実施もある。この場合は、一軸のみ検出するジャイロを一つ用いる。

【0077】次に、利用者位置検出処理34の動作について、図9を用いて説明する。図9は、メモリ131に設ける利用者位置管理テーブル（以下、位置管理テーブルという）36の構成を示す図である。位置管理テーブル36には、利用者位置・動作測定部21で測定した利用者の現在位置を記録する。図9に示す例では、GPS120で測定した経度と緯度と高度を記録している。1

18

8時30分から32分までの間に、利用者は東経140度20分30秒、北緯40度40分35秒、高度10Mの位置から東へ経度6秒分（約180メートル）移動している。図9の例では、経度、緯度は秒単位までであり、高度はメートル単位なので、利用者検出処理の分解能は水平方向で30メートル程度、また垂直方向では1メートルとなっているが、他の応用ではさらに分解能をあげてもよいし、また逆に下げてよい。

【0078】利用者位置管理テーブル36は、利用者が連続して移動する時間（例えば10時間程度）を連続して記録できるように容量を決め、メモリ131に設ける。

【0079】＜メッセージの説明＞次にメッセージ51について、図を用いて説明する。図10はメッセージ51の例を示すものである。この図からわかるように、メッセージが送られるプロセスごとにコマンド58の種類とデータ59の受け渡し方は違う。まず、メッセージID100～102は仮想音空間定義部100（図10ではプロセスIDが0）に送られるメッセージの例である。仮想音空間定義部100では受け取ったメッセージ内の仮想音源IDから処理を施すべき仮想音源を特定し、同じメッセージ内のコマンドを実行する。この際、データ59にはコマンドを実行する際の引数が記述されている。例えば、メッセージID100では、左方向から聞こえてくる仮想音源を右方向へ移動するといった具合に解釈する。仮想音源定義部に対するコマンドは、上述の通りであるが、「強制停止」コマンドでは該当する仮想音源に割り当ててあるプロセスを強制終了する。

【0080】次に、メッセージID200～203ではエージェント群64に対するコマンドとデータをカプセル化している。エージェントプログラムの種類には、例えばPIMエージェント、データベース（以下、DBという）検索エージェント、ニュースウォッチエージェント、位置監視エージェント等がある。PIMエージェントは、例えば利用者のスケジュールや住所録等の利用者個人情報の管理を行う。メッセージID200では7月1日のスケジュールとして関西出張を入力している。この際、利用者が音声でデータを入力する場合は、日付や時間等は音声認識する必要があるが、予定は利用者の音声のままで保存していても構わない。

【0081】メッセージID201ではDB検索エージェントに利用者が現在いる場所の観光名所を探索している。このとき、利用者の現在位置はGPSから得る。ニュースウォッチエージェントは放送されているニュースを利用者の代わりに常にウォッチし、利用者が指定した内容を含むニュースのみ取り出して利用者に聞かせる。このため、利用者はニュースを常に聞いていなくても自分の興味あるニュースを自動的に入手できる。また、この場合はラジオ放送によるニュースだけでなく、テレビ放送によるニュースであっても、エージェントによるウ

(11)

19

ォッチは可能で、放送内容から音声認識し、利用者が指定した話題に関するキーワードとマッチングさせる。

【0082】メッセージID202では米国に関するニュースを集めてくるようにエージェントに命令している。また、この処理結果は、エージェントが最新のニュースを入手する度に利用者にその旨報告する。このとき、エージェントは、処理が終わった時点で、終わったことをまずアラーム、メッセージにて報告し、利用者から改めて指示があつてから結果を報告するようにしてもよい。エージェントからのアナウンスはプロセスが割り当てである左方向から常に聞こえてくるので、利用者は聞こえてくる音声のエージェントからの出力結果であることを判断できる。

【0083】位置監視エージェントは、利用者の現在位置を常に把握し、目的地までの道案内等を行う。もちろん、目的地の位置情報も属性として目的地データの中に備えていて、その相対位置関係を用いる。メッセージID203では東京駅までの道順をナビゲーションするようにエージェントに命令している。この際、現在位置から見た目的地の方向に仮想音源を配置し、利用者が進む方向を直感的に分かるといった工夫もできる。位置監視エージェントが扱う事象としては、道案内の目的地だけではなく、後述する交通渋滞なども可能である。

【0084】次に、メッセージID300～302はハードウェア制御プログラム72に送られるメッセージである。メッセージ300と301では電子会議の際の利用者本人の入力音声メッセージとしてカプセル化され、それぞれプロセスID21と22の電子会議プログラムに送られている。電子会議の相手に送るメッセージが相手毎に分かれているので、それぞれの相手と秘話通話もできる。また、メッセージ302ではラジオ制御プログラムに音量を3レベル落とすよう命令している。

【0085】＜メッセージ作成処理の説明＞次にメッセージ作成処理54について、図を用いて説明する。図11は、選択音源判定処理53を含むメッセージ作成処理54の処理フローを示す図である。流れを説明する前に、選択音源判定処理53（ステップ232、237、243）について説明する。選択音源判定処理53では、利用者が選択（ポインティング）した仮想音源を判定し、メッセージ51内の仮想音源ID56を決める。以下では、本実施例におけるポインティングの方法と選択音源判定処理を場合分けして説明する。マウスに備えられているボタンのように、ポインティングデバイスには選択を確定する手段が必要である。本実施例の利用者が選択する仮想音源の確定について説明する。

【0086】まず、利用者の身ぶりのみで選択を確定する方法について説明する。この場合、あらかじめ選択を確定するジェスチャ（以下、ピッキングジェスチャ：Picking Gestureと呼ぶ）を定義しておく。例えば、利用者が頭部の向きを変えてうなずくと、

20

向いた方向にある仮想音源を選択したとみなす。または、利用者が腕の向きを変えて腕を軽く振ると、腕を向けた方向に配置してある仮想音源を選択したとみなす。この場合、普段の行為と区別するため、2回以上うなずかないと選択が確定されないなど、通常行うことが希なジェスチャにする。

【0087】次に、身ぶりと音声を組み合わせた方法について説明する。先と同じく、選択を確定する音声（以下、ピッキングボイス：Picking Voice）を定義しておく。例えば、「これ」「それ」といった指示語であったり、「選択」といった行為を表す言葉でも構わない。利用者が頭や腕の向きを変えた時に、先のピッキングボイスを発声すると、その頭部向きまたは腕の向きにある仮想音源を選択したとみなす。

【0088】次に、身ぶりと特別なデバイスを組み合わせた方法について説明する。このデバイスとしてはあらかじめ本体部20または利用者動作・位置測定部21に備え付けてあるボタン203、212等のひとつ（以下、ピッキングボタン：Picking Buttonという）を用いる。利用者が頭や腕の向きを変えた時に、ピッキングボタンを押すと、その頭部向きまたは腕の向きにある仮想音源を選択したとみなす。

【0089】また、身ぶりを用いず、すべて音声で行う方法もある（直接指定という）。例えば、「右方向」「左前方」といった具合に、仮想音源が配置してある方向または場所を音声で指定する。あるいは、仮想音源やプロセスに名前等の識別子を付けておいて、その名前を直接呼んで選択する。後者の場合は、仮想音源の名前を指定するのではなく、むしろ「DB検索エージェント」といった具合にプロセス名を直接指定する場合でも、そのプロセスが割り当てである仮想音源をも選択したとみなす。または、仮想音源の位置を画面201または211に表示し、その位置をペン等でクリックするといった仕組みが考えられる。

【0090】同じ方向に複数の仮想音源が配置してある場合は、利用者が指定する方向からだけでは選択された仮想音源を同定することが難しい。本実施例では、利用者がプロセス名で直接指定するか、仮想音源選択後に入力されるコマンドでシステム側が自動的に識別するか、または識別できない場合はシステム側から利用者に問い合わせるといった方法をとる。上記いずれの選択方法においても、選択された仮想音源を利用者に確認させるという意味で、仮想音源から出力される音量を大きくしたりして変化させるといったことも有効である。

【0091】次に全体の流れについて説明する。まず、ステップ230では利用者操作を検出する。利用者が行う操作には、ジェスチャによる操作と、音声による操作と、ボタンによる操作の三種類がある。ジェスチャを検出する場合は、常に利用者動作管理テーブル35（図7参照）を監視しておき、頭部または手の向きの変化量が

(12)

21

あるしきい値を越えた場合に利用者がジェスチャを入力したとみなす。また、音声が入力された場合とボタンが押された場合は、それぞれ音声による操作とボタンによる操作とみなす。ステップ 230 でジェスチャ、音声またはボタンの入力を検知されない場合はなにもせずにメッセージ作成処理 54 を抜ける（ステップ 248）。メッセージ作成処理 54 はタイマを用いてある一定時間ごとに起動するようにし、これによって利用者操作検出を繰り返す。

【0092】ステップ 231 ではそのジェスチャを解析し、それがピッキングジェスチャまたはコマンドジェスチャであるかを判定する。ここで、コマンドジェスチャとは利用者がコマンドを入力するためのジェスチャを意味し、例えば頭部を横方向に振ると選択済み仮想音源の音出力を一時停止するといった具合に、ジェスチャとコマンドの対応をあらかじめ決めておく。そして、ピッキングジェスチャとコマンドジェスチャともに利用者の標準ジェスチャをあらかじめ登録しておき、入力ジェスチャとのマッチングをとる。利用者は常に同じ動作を忠実に繰り返すという前提は現実的でなく、特に時間軸のズレは大きいと考えられる。そこで、ジェスチャの特異変化点を符号化してその符号列でマッチングする方法や、時間軸のズレを吸収する DP マッチングアルゴリズムを使って標準ジェスチャと入力ジェスチャの変化パターンを比較する方法を用いる。

【0093】ピッキングジェスチャを検出すると、現在の頭部向きまたは手の向きにある仮想音源を選択したとみなし、仮想音源 ID をメッセージ内にセットする（ステップ 232）。

【0094】コマンドジェスチャを検出した場合は、そのコマンドがどのプロセスに対するコマンドであるかをまず識別する（ステップ 233）。例えば「移動」、「一時停止」、「出力優先」等の仮想音源に対するコマンドを検出した場合は、メッセージ転送先のプロセスを仮想音空間定義部にセットする。それ以外のコマンド（つまりユーザプロセスに対するコマンド）を検出した場合は、既にメッセージ内にセットしてある仮想音源 ID を調べて、その仮想音源に割り当ててあるプロセスを転送先にセットする。転送先としては、エージェントやハードウェア制御プログラムになる。この際、プロセス ID は、図 12 に示す仮想音空間管理テーブルから選択済み仮想音源 ID に該当するものを検索する。また、検出したコマンドが当該プロセスに合致しない場合は、ステップ 235 で利用者に警告を発し、利用者操作検出に戻るか、または、利用者に確認をうながす。

【0095】ステップ 234 では、コマンド内容とデータをメッセージにセットする。ここで、メッセージが完成したことになるので、メッセージ転送処理 55 によって適切なプロセスにメッセージが送られる。

【0096】入力ジェスチャがピッキングジェスチャと

22

コマンドジェスチャのいずれとも違う場合は何もせずにメッセージ作成処理を抜け、再度利用者操作検出を繰り返す（ステップ 235）。

【0097】同じく音声が入力された場合は、ステップ 236 で入力音声を認識し、それがピッキングボイスまたはコマンドボイスであるかを判定する。ここで、コマンドボイスとは利用者が音声で入力したコマンドを意味する。先にも述べたように、音声認識には特定話者の単語音声認識を用いるので、あらかじめピッキングボイスとコマンドボイスともに利用者の標準音声を登録しておき、入力音声とのマッチングをとる。

【0098】ピッキングボイスには、身ぶりや手振りによる選択を確定するための「これ」「それ」といった指示語や「選択」といった行為を表す言葉をあらかじめ設定しておく。これらピッキングボイスを検出した際は、その時点における頭部向きまたは手の向きにある仮想音源を選択したとみなし、仮想音源 ID をメッセージ内にセットする（ステップ 237）。

【0099】「DB 検索エージェント」といった具合に、プロセス名を音声で直接指定する場合は、プロセス名で指定したプロセス ID とそのプロセスが配置してある仮想音源の ID をメッセージ内にセットする（ステップ 238）。この場合も、図 12 に示す仮想音空間管理テーブルを参照し、直接指定したプロセスに対応する仮想音源 ID を求める。

【0100】コマンドボイスを検出した場合の処理（ステップ 239 及びステップ 240）は、コマンドを音声認識によって識別すること以外は、コマンドジェスチャを検出した場合のステップ 233 とステップ 234 とほとんど同じなので詳細は省く。同じく、検出したコマンドが選択済みの仮想音源に割り当ててあるプロセスと矛盾する場合や入力音声ピッキングボイスとコマンドボイスのいずれとも合致しない場合は、ステップ 235 と同様に利用者に警告を発し、利用者操作検出に戻る（ステップ 241）かまたは利用者に確認をうながす。

【0101】ボタンによる入力検出された場合は、ステップ 242 で利用者が選択したボタンを識別し、それがピッキングボタンまたはコマンドボタンのいずれかであるかを判定する。ここで、コマンドボタンとは利用者がボタンで入力したコマンドを意味する。

【0102】ピッキングボタンには、身ぶりや手振りによる選択を確定するためのボタンをあらかじめ決めておく。このピッキングボタンを検出した際は、その時点における頭部向きまたは手の向きにある仮想音源を選択したとみなし、仮想音源 ID をメッセージ内にセットする（ステップ 243）。

【0103】プロセス名を表示したボタンを直接指定する場合は、プロセス名で指定したプロセス ID とそのプロセスが配置してある仮想音源の ID をメッセージ内にセットする（ステップ 244）。ステップ 238 と同じ

(13)

23

く、仮想音空間管理テーブルを参照し、直接指定したプロセスに対応する仮想音源IDを求める。

【0104】コマンドボタンを検出した場合の処理（ステップ245及びステップ246）は、コマンドをボタン選択／非選択によって識別すること以外は、コマンドジェスチャを検出した場合のステップ233とステップ234とほとんど同じなので詳細は省く。検出したコマンドが選択済みの仮想音源に割り当ててあるプロセスと矛盾する場合は、利用者に警告を発し、利用者操作検出に戻るか、または、利用者に確認をうながす。また、その他のボタンが押された場合は、そのボタンに割り当ててある処理を実行する（ステップ247）。

【0105】次に、仮想音空間定義部100と、その中の仮想音源位置設定処理104について、図12を用いて詳しく説明する。図12に示す出力状態は、各仮想音源からの音出力の状態を表し、通常出力と最小出力と優先出力と一時停止等の状態をとる。直感的には、仮想音源からの出力音の音量であらかじめ決められたレベルととらえても構わない。つまり、通常出力の場合は、他の仮想音源からの出力音と同等に扱われ、すべて同程度の音量レベルで出力される。最小出力では、最小限聞こえる程度の音量レベルで音が出力される。優先出力では、他の仮想音源からの出力音はすべて最小出力状態になり、結果的に当該仮想音源からの出力音が優先されることになる。一時停止状態では、当該仮想音源の出力音量レベルが0になり、利用者には何も聞こえなくなる。

【0106】仮想音源位置は、感覚上の音源が利用者からみてどこに配置されているかを表す。図12中の水平位置Xと水平位置Yと高さZは、利用者の両耳を結ぶ線をX軸にとり、両耳間の中心を通り当該X軸と垂直をなす線をY軸に取った座標上での交点を原点としたときのそれぞれX座標とY座標を表す。同様に、高さZは利用者の耳の位置からみた高さ方向位置を表す。図12中ではともにメートル単位で記入してある。このX軸、Y軸は、前述の水平方向の基準面221に含まれる。さらに、前述の利用者動作管理テーブル35に記録する回転角は、このX軸、Y軸の交点（原点）を回転中心とする。

【0107】メッセージ作成処理54が作成したメッセージ51を、その内容に従って、メッセージ転送処理55が該当するユーザプロセスに送ると、仮想音源位置設定処理104が当該ユーザプロセスに適当な仮想音源を割り当てて、さらに当該仮想音源をデフォルトの位置に配置するように、仮想音源位置管理テーブル106を設定する。

【0108】いったん仮想音源の位置が決まると、あとは利用者が自由に配置を換えたり、各仮想音源からの出力音の音量を調整できる。利用者は先にも述べた各コマンドを入力し、仮想音源の配置や出力音量を調整する。これらのコマンドは仮想音源IDとともにメッセージと

24

してカプセル化され、仮想音空間定義部100に送られてくる。

【0109】仮想音空間定義部100では、メッセージ内の仮想音源IDとコマンドをみて仮想音空間管理テーブル106を修正する。ここで、「移動」コマンドを受け取った場合は、該当する仮想音源IDの音源位置をメッセージ内のデータに従って書き換える。また、「音量最小化」、「音量最小化解除」、「一時停止」、「一時停止解除」、「出力優先」、「出力優先解除」の各コマンドを受け取った場合は、テーブル内の出力状態を書き換える。この際、「出力優先」の場合は指定された仮想音源（または仮想音源群）以外はすべて自動的に最小出力状態にし、「出力優先解除」とともに元の状態に戻す。「音量変更」では各音源からの出力音量をアナログ的に調整する。

【0110】また、人間の聴覚は方向性に比べて距離感の音像定位は鈍いので、仮想音空間で利用者と仮想音源間の距離を変化させることは音量を変化させることで対応する。最後に、「強制停止」コマンドでは、メッセージで指定された仮想音源IDのデータを仮想音空間管理テーブル106から削除し、同時に当該仮想音源に割り当ててあったプロセスをメモリ131内から削除する。

【0111】また、システム側で自動的に仮想音源の位置や出力音量を調整する場合もある。特に、緊急を要する情報や利用者があらかじめ命じておいた処理の結果はシステム側で自動的に優先出力する必要がある。このためには、図13に示すように、ユーザプロセスが出力する音情報に重要度や内容識別子といった音情報属性を付加できるようにしておき、同時に仮想音空間定義部内の属性に対する音源位置決め規則も変更する。例えば、音声メールを相手利用者に送る場合を想定すると、送る先に早急に聞いて欲しい内容は重要度を高く設定する。

【0112】仮想音空間定義部では、当該音情報が他の音情報に優先して出力されるように加工し、受け手側では当該音情報が他の音情報に優先して出力される。また同様に、定義しておくことで、相手側でも即座に聞きたい内容をあらかじめメールエージェントに伝えておき、当該メールエージェントは音情報属性内の内容識別子から送られてきたメールの内容を判断し、送り先が設定した重要度にかかわらず当該音情報を優先出力する。

【0113】この際、優先出力する前にエージェントから利用者にその旨伝達し、利用者からの許可を待って優先出力する仕組みにする。なお、他のメールについては利用者が好きなときにいつでも聞けるものとする。また、ニュースウォッチエージェントが利用者が要求する内容に合致したニュースを入手した場合も同様にシステム側で優先出力する。

【0114】図12に示すような仮想音空間が定義されると、続いて仮想音空間作成部110は、この定義の通りに各音情報が利用者に聞こえるような仮想音空間を作

(14)

25

成する。具体的には、音情報内の仮想音源 I D (図 1 3 参照) をキーとして、仮想音空間管理テーブル 1 0 6 から当該音情報の仮想音空間における状態を検出し、その状態にあうよう音データを加工する。各音データが加工されると、それらをヘッドホンまたはスピーカの配置にあうように合成して出力する。音像定位加工及び合成のやり方で大きく二種類の構成が考えられ、それぞれ図 1 4 と図 1 5 に構成例を示す。

【0 1 1 5】図 1 4 では、仮想音空間内に定位する音像の数と位置があらかじめ固定してあり、入力される音情報の数が、その音像の数と一致またはそれ以下である場合の構成を表し、音入力 2 5 0 ~ 2 5 1 と仮想音空間管理テーブル 2 5 2 とクロスバースイッチ 2 5 3 と音出力 2 5 4 ~ 2 5 5 からなる。クロスバースイッチ 2 5 2 では、仮想音空間内での音像位置に対応するように、音入力と音出力を対応付ける。

【0 1 1 6】具体的には、音入力内にどの仮想音源から出力すべきかを記載しておき、また仮想音空間管理テーブル 2 5 2 には各仮想音源がどの音出力に割り付けてあるかを記載しておく。クロスバースイッチ 2 5 3 は当該仮想音空間管理テーブルを通じて音入力と音出力を対応付ける。例えば、音出力手段としてヘッドホンを用いる場合を考えると、入力音は必然的に二種類に限られ、クロスバースイッチ 2 5 3 が入力音をヘッドホンの右及び左の出力に対応付ける。この場合は構成が簡単で、安価に実施できる。

【0 1 1 7】図 1 5 は、音像の数を音出力手段の数に依存させない場合の仮想音空間作成部の構成を示すものである。。音入力 2 6 0 ~ 2 6 1 から音出力 2 7 0 ~ 2 7 1 にマッピングするフィルタ 2 6 4 ~ 2 6 7 を設け、この結果を各出力手段ごとに設けたミキサで合成する。ここで音像の数を可変にできるよう、フィルタやミキサにはプログラム可能な DSP (Digital Signal Processor) 等を用いる。また、音出力手段がヘッドホンなどのように音信号を両耳に直接伝達する場合とスピーカによっていったん音響空間を経て利用者に伝達する場合の二種類について処理方法を考える必要がある。以下それぞれについて説明する。

【0 1 1 8】まず、ヘッドホンを用いる場合には、バイノーラル方式と呼ばれる音像定位方法を用いる。音源を置いた部屋でダミーヘッドの両耳にマイクロホンを取り付けて録音したものをヘッドホンを用いて再生すると、録音した部屋と同じ音場を体験できる。この原理を応用し、各音源の位置から両耳までの音響的な伝達関数を求め、各音信号にこの伝達関数を畳み込み、両耳ごとに積分することにより、仮想音空間を作成する。この伝達関数は、無響室などにおいて音源を部屋内で移動させ、各地点でダミーヘッドの両耳に取り付けたマイクロホンから音を録音して測定する。伝達関数には個人差があるので、複数の代表的な伝達関数を用意し、利用者にあらか

26

じめ自分にあった伝達関数を選んでもらってもよい。

【0 1 1 9】次に、音出力手段がスピーカである場合は、ヘッドホンを用いる場合に比べて、スピーカから両耳に至るまでの空間を音信号が伝達する際の影響を新たに考慮する必要がある。ここで、利用者の位置は固定しているとすると、スピーカから利用者の両耳までの伝達関数を求め、これを先のバイノーラル方式で求めた伝達関数と組み合わせたフィルタを用いればよい。本方式は、トランスノーラル方式と呼ばれる。スピーカから見た利用者の位置は固定していることを前提とするので、車載型情報通信システム等で本発明を実施する場合に用いる。また、音出力手段は 2 個より多ければ、より多くのスピーカで仮想音空間を作成できるので、安定した音像定位が得られる。

【0 1 2 0】図 1 5 の音入力 2 6 0 には仮想音源 2 7 2 が対応している。これは、音情報内の仮想音源 I D から判断する。また、当該仮想音源 2 7 2 の仮想音空間内での位置は仮想音空間管理テーブル 1 0 6 からわかる。その音源位置から音出力手段 2 7 0 を通じて利用者の耳に伝達されるまでの伝達関数がフィルタ 2 6 4 に対応する。一方、フィルタ 2 6 5 は音出力手段 2 7 1 を通した場合の伝達関数であり、フィルタ 2 6 6 と 2 6 7 は音入力 2 6 1 に対応する仮想音源が置かれた位置から出力手段 2 7 0 と 2 7 1 を通した場合の伝達関数に相当する。それぞれ入力された音を伝達関数で畳み込み、これを各音出力手段に対応して設けられたミキサ 2 6 8 ~ 2 6 9 で合成する。なお、図 4 中の音像定位加工処理 1 1 1 は図 1 5 中の処理 2 6 2 ~ 2 6 3 に対応する。同じく、音合成処理 1 1 2 はミキサ 2 6 8 ~ 2 6 9 での処理に対応する。

【0 1 2 1】フィルタ 2 6 4 ~ 2 6 7 の伝達係数は、各仮想音源位置と各出力手段とのマトリックス (図 1 6 参照) としてディスク 1 3 2 中に保存されている。なお、図 1 6 では、簡単のため、仮想音源位置は四種類しか記載されていないが、もちろん多様な仮想音源位置を設定することができる。この場合、人間が音像を聞き分ける能力とシステムによる音像定位の性能とを考慮に入れて仮想音源位置の分解能を決める。

【0 1 2 2】＜ソフトウェア全体構成の説明＞図 1 7 に、これまでに述べたソフトウェア構成図の全体の大まかな流れを示す。ステップ 2 8 3 からステップ 2 8 8 はすでに述べたので、ここではステップ 2 8 1 とステップ 2 8 2 の前処理と、ステップ 2 8 9 の後処理について述べる。プログラム起動時には仮想音空間をデフォルトの状態に設定する (ステップ 2 8 1)。利用者が前回終了時の状態で起動したい場合は、その旨設定できるように、ステップ 2 8 9 にて常に終了時の仮想音空間の状態をディスク 1 3 2 中に保存しておく。また、音情報は視覚情報と比べて一覧性に劣るので、どのユーザプロセスが仮想音空間中のどの位置に割り当てられたかを利用者

(15)

27

に知らせる必要がある。ステップ282では、例えば、各ユーザプロセスごとにプロセス名と現在位置を利用者にアナウンスすることによって、仮想音空間の初期状態を知らせる。また、利用者は起動後も任意に仮想音空間の状態を知ることができる。

【0123】＜実施例の利用例＞以下、図を用いて上記実施例における音情報提供装置の利用例を示す。図18は利用者による音情報の選択及び入力とその結果出力される音情報との関連を時間軸に従って記載したものである。図19は図18の時刻t1からt2までの仮想音空間290の状態を表したものである。ラジオからの出力音は音像292に、B氏との電子会議からの出力音声は音像293に、同じくA氏からの出力音声は音像294に、電子秘書からの合成音声の出力は音像295に配置してある。

【0124】まず、時刻t1までは右方向の音情報としてラジオからの出力が選択されている。時刻t1で電子会議を行うための呼出音が前方から聞こえ、この前方の音を利用者が選択することにより電子会議が開始される。選択方法は前述のいずれでもよい。利用者動作検出処理33は、ジャイロ121、122の動作を図7の動作管理テーブル35に記録し、メッセージ作成処理54が、図11に示すステップに従って、選択内容を判定してメッセージ51を作成し、メッセージ転送処理55がハードウェア制御プログラム群72にメッセージを転送する。

【0125】ハードウェア制御プログラム群72はPHS通信機163、164を介した電子会議相手の声を音情報115として渡す。この時、会議相手の配置はあらかじめ仮想音空間定義部100を介して定義しており、仮想音空間作成部110によって、A氏の音声は左前方から聞こえ、B氏の音声は右前方から聞こえる。

【0126】時刻t2で、利用者は電子会議で必要となった情報をデータベースから検索するため、あらかじめ左方向に配置しておいた電子秘書を起動する。利用者291が方向299と方向300の間を向いて、選択動作を行うと、利用者状態検出部30によってコマンド46が作成される。コマンド46は、メッセージ転送部50によってメッセージ51となり、エージェント群64に渡される。この結果、エージェント群64のうち、たとえば、DB検索エージェントが起動される。このとき、仮想音空間作成部110は、電子秘書の音像295が仮想音空間内を利用者291に近づく方向に移動し、電子会議の音像293と294がともに利用者291から離れる方向に移動するように、音像を制御する。なお、上述のように、仮想音空間で利用者との距離を変化させることは音量を変化させることで対応する。

【0127】電子秘書（エージェント群64）が起動されると、利用者は、データベース検索命令と検索条件を入力する。検索内容をマイク23から入力すると、入力音声データ43は、音声認識部41で認識され、コマン

28

ド44、データ45が作成される。コマンド44、データ45は、メッセージ転送部50によってメッセージ51となり、エージェント群64に渡される。

【0128】エージェント群64がデータベースを検索している時刻t3から時刻t4の間、利用者は前方の音情報を選びなおし、電子会議を再開する。図18中の時刻t3で前方を指し示すと、上述と同様の過程を経て、再度電子会議が選択される。ここで、利用者が指し示した方向は298であり、音像293と音像294との中間にあたる。これは、A氏及びB氏両者との電子会議を選択したことを意味する。

【0129】時刻t4で、検索処理が終わると、エージェント群64は左方向から、音声合成部63で合成した音声またはアラーム音で知らせる。利用者が電子秘書を再度選ぶと、検索結果を音声合成部63で音声合成し、音情報115として出力する。この時、出力内容は同時に音情報一時記憶領域114に蓄積される。利用者は聞き逃したり即座に理解できなかった場合に、入力音声データ43、コマンド46を入力することで、何度も繰り返し聞き直すことができる。

【0130】時刻t5で、利用者は、検索結果についてA氏のみと話し合うため、左前方の音情報を選択する。選択は上述のように行う。時刻t5では方向298と方向299の丁度真ん中あたりにある音像294を指し示したため、メッセージ作成処理33はA氏のみと電子会議を再開するようメッセージ51を作成し、ハードウェア制御プログラム群72に転送する。人間の聴覚は左右水平方向の音源には敏感で、2°～3°程度の誤差で音源の場所を言い当てることができる。なお、この時、右前方からB氏からの音声も聞こえていて構わないが、B氏には利用者からの音声は届かないように、ハードウェア制御プログラム群72はPHS通信機63または164を制御する。これは電子会議で特定の相手と「ひそひそ話し」する場合などに適用できる。

【0131】時刻t6では、B氏もまじえて電子会議を再開し、時刻t7で電子会議を終了する。終了するとともに、これまで電子会議開始後から一時停止状態だったラジオからの出力が自動的に再開される。

【0132】仮想音空間の状態を図5に示す画面201または図6に示す画面211に図示して、ボタンで選択するようにしてもよい。また、利用者操作により、ラジオの音像292と電子秘書の音像295を入れ替えるなど、音像の位置を変更することも可能である。本実施例では、GPS（または室内位置検出用の赤外線データ受信機）が腕時計型の本装置21に内蔵されると述べてきたが、必ずしもこのような構成である必要はなく、例えばGPSが本体部20に内蔵されていても構わない。

【0133】また、利用者の手の動きを測定する手段として、ジャイロの代わりにデータグローブを用いる構成も考えられる。データグローブは、グローブの各指の上

(16)

29

部に光ファイバが通してあり、指を伸ばしたり曲げたりする際の光の屈折率によって、手の動きと現在の手の形状を検出する。従ってデータグローブを装着した利用者の手や指の動きも容易に検出できる。さらには、利用者の指輪や腕時計に発光体を装着し、それらをはめた手をビデオカメラで撮影するといった構成も考えられる。ここでは、手の映像を画像処理し、発光体の位置を抽出することによって利用者の手の動きを測定する。

【0134】<他の実施例>つぎに、他の実施例について図を用いて説明する。図20と図21は、本発明の実施例に係る音情報提供装置の他の利用場面及び装置構成を表した図である。以下、それぞれ説明する。

【0135】図20は、本発明を車載型情報通信システムで用いた場合の利用例を表す。最近の自動車310では、自動車電話313や、CDプレーヤ314や、カーラジオ・TV315等が備え付けられたものが多い。さらには、自動車ナビゲーションシステムなどのように、自動車測位装置316も備え付けているものもある。一般に、これは別々に制御され、スピーカ319～320（前述のスピーカ151に対応する）から出力される。ここで、本発明による音情報提供装置311と、利用者の音声を入力するマイク312と、利用者の動作を検出するためのビデオカメラ317及び指輪型の発光体318を新たに設置する。

【0136】この実施例では、利用者の現在位置に応じた情報を提供する。例えば、高速道路を走りながら音楽を聞いていて、あるポイントに車が近づくと進行方向から自然に道路渋滞情報が聞こえてくる。提供される情報にはそれが利用者にとって有効と思われる有効範囲を示すデータも含まれており、利用者がその場所に近づいて初めて情報が自然に聞こえてくる。また、その情報に関連する事象が発生した場所に関するデータも位置属性として提供され、聞こえてくる方向は利用者の現在位置から見た情報発生場所の方向と対応付けられている。道路渋滞情報の場合、渋滞している方向から音情報が聞こえると、利用者は不慣れな場所を運転していても、どこが渋滞しているかを容易に判断できる。

【0137】これを実現する際は、まず前記GPSなどの利用者状態検出部30から利用者の現在位置を入手し、提供される情報に含まれる有効範囲のデータと比較し、現在位置が有効範囲内にある場合は当該音情報を出力する。このことにより、先に説明したような利用者の現在位置に応じた道路交通情報の入手など、今までにない新しい効果を得ることができる。

【0138】これを発展させ、利用者が指し示した方向から提供される情報を出力するなどの応用も考えられる。例えば、電車の中吊り広告に音情報の発信器が取り付けられてあり、音情報が放送されている。利用者がその中吊り広告を指し示したことによって、その中吊り広告が発信している音情報を選択したことになり、利用者には

30

指し示した方向から当該音情報が聞こえてくる。

【0139】なお、第1の実施例において、利用者の動作は秒単位で変化するが、位置の検出では本実施例の最小単位が30メートルであり、利用者が歩いて移動することを前提とすると利用者位置は急激に変化しない。そのため、図7と図9で時刻の分解能を変えていた。しかし、本実施例において、本装置を車載システムとして実施する場合は、図9の時刻の分解能はさらに細かくとるべきである。また、歩行と乗り物による移動が併用される場合は、位置の変化量から時刻の分解能を動的に変えるといった方式をとる。例えば、利用者位置を検出する時間間隔の上限値を設定して、その範囲内で前時刻から現時刻までの利用者位置の変化量に反比例させて時間間隔を決める。

【0140】最後に、図21は本発明による音情報提供装置を室内で利用する場合の応用例を表す。室内330の天井には、利用者からのデータを送受信する赤外線データ送受信機332と、スピーカ334～335と、利用者が部屋にいるかどうかを識別するためのビデオカメラ333が備え付けてある。利用者は腕時計型の赤外線データ送受信機331を内蔵した利用者位置・動作測定部21を携帯している。前記複数の赤外線データ送信機332は、たえず送信機の位置と時刻のデータを送信する。そして、利用者が身に付けている赤外線データ受信機331は、前記送信機からのデータを受信して、各送信機からの距離を計算する。計算した利用者の現在位置は、常に赤外線データ送受信機332に送信している。

【0141】また、本装置を室内などある決まった領域で用いる場合は、緯度、経度、高度に相当する何らかの基準を決めておく必要がある。例えば、基準となる原点と水平方向のXY軸を適当に決めて、原点から見た利用者の現在位置をメートル単位で利用者位置管理テーブル36に記録する。

【0142】サーバコンピュータ61は、利用者の現在位置を常に把握し、その場所に応じた音情報を出力する。例えば、室内のある領域が立入制限区域になっており、許可を受けていない利用者がその領域に近づくと、制限区域の方向から警告メッセージが流れる。利用者は警告を知るだけでなく、警告メッセージが聞こえる方向からどこが立入制限区域であるかも判断できる。また、本応用例では利用者が動きまわることを前提としているため、音像定位の方法として、図14による方法を用いるが、この時スピーカ334～335はコンピュータ制御により自由に配置できることが望ましい。図21ではスピーカ334～335を天井に配置しているが、スピーカを床に配置し自由に移動できる手段を設けた応用例も考えられる。

【0143】また、本発明は視覚情報を用いずに情報提供できるという点で、目の不自由な障害者向けの情報提供装置としても応用できる。その場合は、本体部20と

(17)

31

利用者位置・動作測定部21は、ボタン203、212がより重要になり、画面201、211に表示されるソフトウェアボタンよりは、各キーの形状、キートップの凹凸に特徴を持たせた機械的なボタンの方が望ましい。

【0144】

【発明の効果】以上述べたように、本発明によれば、複数の音情報を音像定位し、合成することにより、利用者は複数の音が同時にそれぞれの方向から聞こえてくるように感じることができる。

【0145】さらに、利用者は同時に聞こえる複数の音情報の中から優先して聞きたい音情報を選択したり、聞きたくない音情報を消したり、音情報の聞こえる方向やそれぞれの音量を自由に変更したりできる。

【0146】さらに、利用者が選択した音情報に割り当てられた処理に利用者の命令またはデータを送ることができる。

【0147】以上により、音情報における「ウィンドウ」や「マウス」に相当する概念を提供することができる。この結果、利用者は仮想的な音空間を渡り歩くことによって平行して仕事を進めることができ、人間とコンピュータとの対話及び人間と人間との対話の効率も向上する。

【図面の簡単な説明】

【図1】本発明の特徴を概念的に表す説明図である。

【図2】本発明の一実施例に係る音情報提供装置の利用イメージと基本構成を表す説明図である。

【図3】本発明の一実施例に係る音情報提供装置のハードウェア構成を表す説明図である。

【図4】本発明の一実施例に係る音情報提供装置のソフトウェア構成を表す説明図である。

【図5】本体部の装置外観を表す説明図である。

【図6】利用者位置・動作測定部の装置外観を表す説明図である。

【図7】利用者動作管理テーブルの構成を表す説明図である。

【図8】利用者頭部向きを検出する場合の例を表す説明図である。

【図9】利用者位置管理テーブルの構成を表す説明図である。

【図10】メッセージの例を表す説明図である。

【図11】メッセージ作成処理の処理フローを表す説明図である。

【図12】仮想音空間管理テーブルの構成を表す説明図である。

【図13】音情報の構成を表す説明図である。

【図14】仮想音空間作成部の構成例を表す説明図である。

【図15】仮想音空間作成部の別の構成例を表す説明図である。

【図16】図15における仮想音空間作成部のフィルタ

32

を管理するテーブルを表す説明図である。

【図17】ソフトウェア構成図の全体の処理フローを表す説明図である。

【図18】本発明の一実施例に係る音情報提供装置の利用例を表す説明図である。

【図19】図18の時刻t1からt2までの仮想音空間の状態を表す説明図である。

【図20】本発明を車載型情報通信システムで用いた場合の応用例を表す説明図である。

【図21】本発明による音情報提供装置を室内で利用する場合の応用例を表す説明図である。

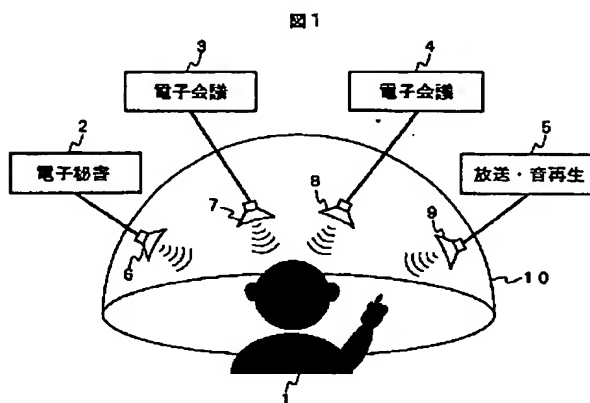
【符号の説明】

1…利用者、	2…電子秘書、
3…電子会議、	4…電子会議、
5…放送・音再生、	6～9…音像、
10…仮想音空間、	20…本体部、
21…利用者位置・動作測定部、	22…利用者頭部向き測定部、
23…マイク、	
24…ヘッドホン、	30…利用者状態検出部、
31…利用者動作、	32…利用者位置、
33…利用者動作検出処理、	34…利用者位置検出処理、
35…利用者動作管理テーブル、	
36…利用者位置管理テーブル、	40…音声入力部、
41…音声認識部、	42…入力部、
43…入力音声データ、	44…コマンド、
45…データ、	46…コマンド、
47…データ、	50…メッセージ転送部、
51…メッセージ、	
53…選択音源判定処理、	54…メッセージ作成処理、
55…メッセージ転送処理、	56…仮想音源ID、
57…プロセスID、	58…コマンド、
59…データ、	60…電子秘書、
61…サーバ、	62…データ通信手段、
63…音声合成部、	70…放送受信手段、
71…放送局、	72…ハードウェア制御プログラム群、
80…音再生手段、	81…CD、
90…音声通信手段、	91…音声通信手段、
92…通信相手、	93…通信相手、
100…仮想音空間定義部、	
101…仮想利用者位置、	102…仮想音源位置、
103…音情報属性、	104…仮想音源位置設定処理、
105…仮想利用者位置設定処理、	106…仮想音空間管理テーブル、
110…仮想音空間作成部、	111…音像定位加工処理、
112…音合成処理、	113…仮想音空間出力部、
114…音情報一時記憶領域、	115…音情報、
120…GPS、	121…ジャイロ、
122…ジャイロ、	130…CPU、
131…メモリ、	13

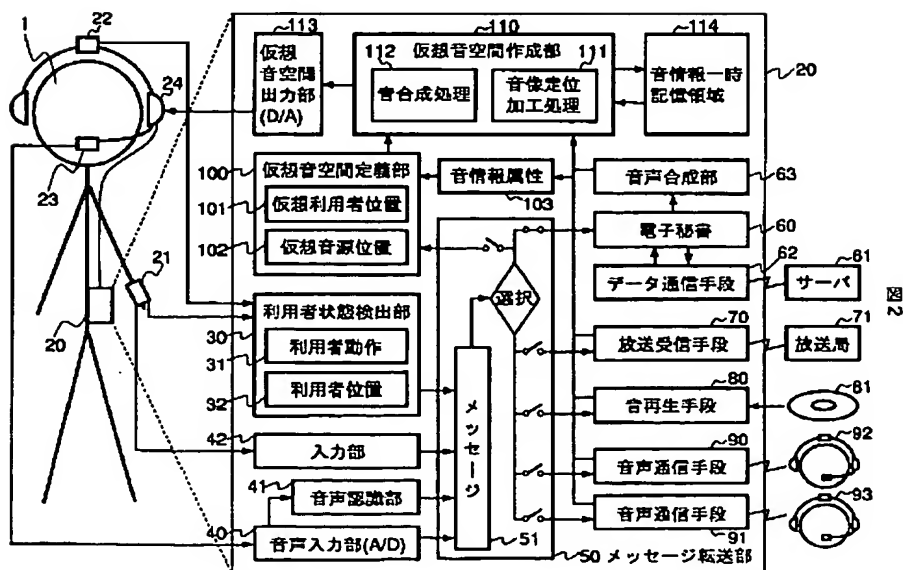
(18)

33
 2…ディスク、
 換器、141…ボタン類、
 …D/A変換器、151…スピーカ、
 160…赤外線通信機、161…ラジオ受信機、
 162…CDプレーヤ、163、164
 …PHS通信機、165～168…A/D変換
 器、169…バス、201…
 液晶画面、202…ペン、
 03…ボタン、204～206…PCカード用スロ
 ット、207…赤外線通信機、208…ファックスモデ
 ム、209…カバー、211…液晶画面、
 212…ボタン、213…マイ
 ク、214…バンド、250～
 251…音入力、252…仮想音空間管理テーブル、
 253…クロスバースイッチ、254～255…音

【図1】



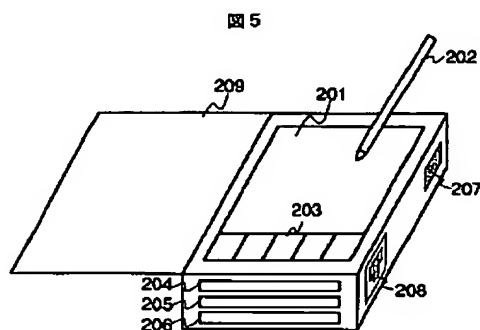
【図2】



34

出力、260～261…音入力、264
 ～267…フィルタ、268～269…ミ
 キサ、270～271…音出力、272
 ～273…仮想音源、291…利用者、
 292～295…音像、296～300…方
 向、310…自動車、311…音情報
 提供装置、312…マイク、313…自
 動車電話、314…CDプレーヤ、
 315…カーラジオ・TV、316…自動
 車測位装置、317…カメラ、
 318…発光体、319～320…スピーカ、
 330…室内、331…赤外線データ送受信機（携
 帯型）、332…赤外線データ送受信機（据置型）、3
 33…カメラ、334～335
 …スピーカ

【図5】

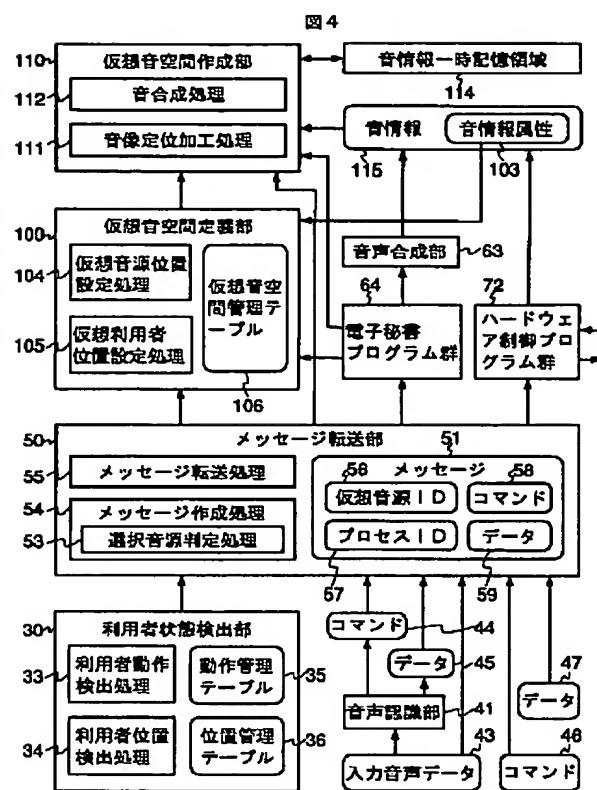


(19)

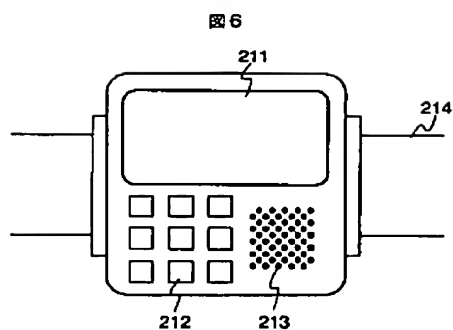
【図3】



【図4】



【図6】



【図7】

図7 35 利用者動作管理テーブル

時刻	手の向き (水平方向、垂直方向)	頭の向き (水平方向、垂直方向)
18:30:40	(0°、 0°)	(0°、 0°)
18:30:41	(0°、 0°)	(-25°、 0°)
18:30:42	(0°、 0°)	(-45°、 0°)
...

【図9】

【図8】

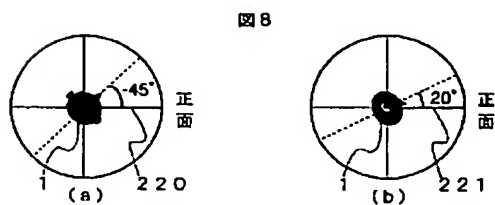


図9 36 利用者位置管理テーブル

時刻	経度	緯度	高度
18:30	東経140度20分30秒	北緯40度40分35秒	10M
18:31	東経140度20分33秒	北緯40度40分35秒	10M
18:32	東経140度20分36秒	北緯40度40分35秒	10M
...

(20)

【図10】

メッセージID	仮想音源ID (位置)	プロセスID (種類)	コマンド	データ
100	11 (左)	0 (仮想音空間定義部)	移動	右へ
101	21 (左前方)	0 (仮想音空間定義部)	一時停止	—
102	31 (右前方)	0 (仮想音空間定義部)	出力優先	—
200	11 (左)	10 (PMI-ジェント)	登録	7月1日は関西出張予定
201	12 (左)	11 (DB検索エージェント)	検索	この地域の観光名所
202	13 (左)	12 (ニュースエージェント)	ウォッチ	米国に関するニュース
203	14 (左)	13 (位置監視エージェント)	ナビゲーション	東京駅までの道順
300	21 (左前方)	21 (電子会議プログラム)	接続	(入力音声データ)
301	31 (右前方)	22 (電子会議プログラム)	接続	(入力音声データ)
302	41 (右)	23 (ラジオ制御プログラム)	音量変更	—3レベル

図10

【図12】

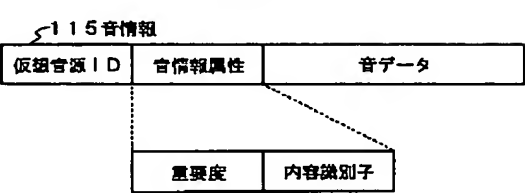
図12

106仮想音空間管理テーブル

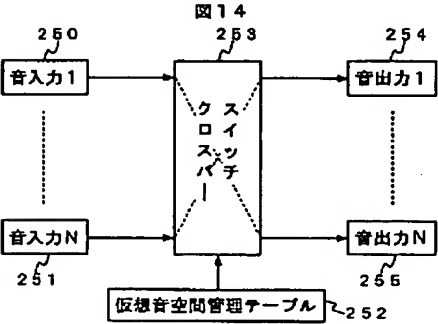
仮想音源ID	プロセスID	出力状態	水平位置X	水平位置Y	高さZ
11	11	一時停止	—3M	0M	0M
21	21	優先出力	—1M	1.5M	0M
31	22	優先出力	1M	1.5M	0M
41	23	一時停止	3M	0M	0M
...

【図13】

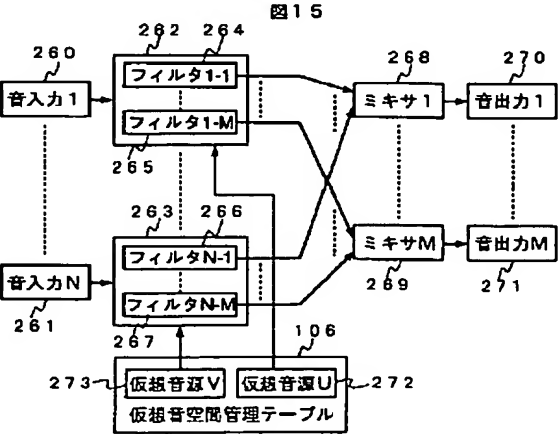
図13



【図14】



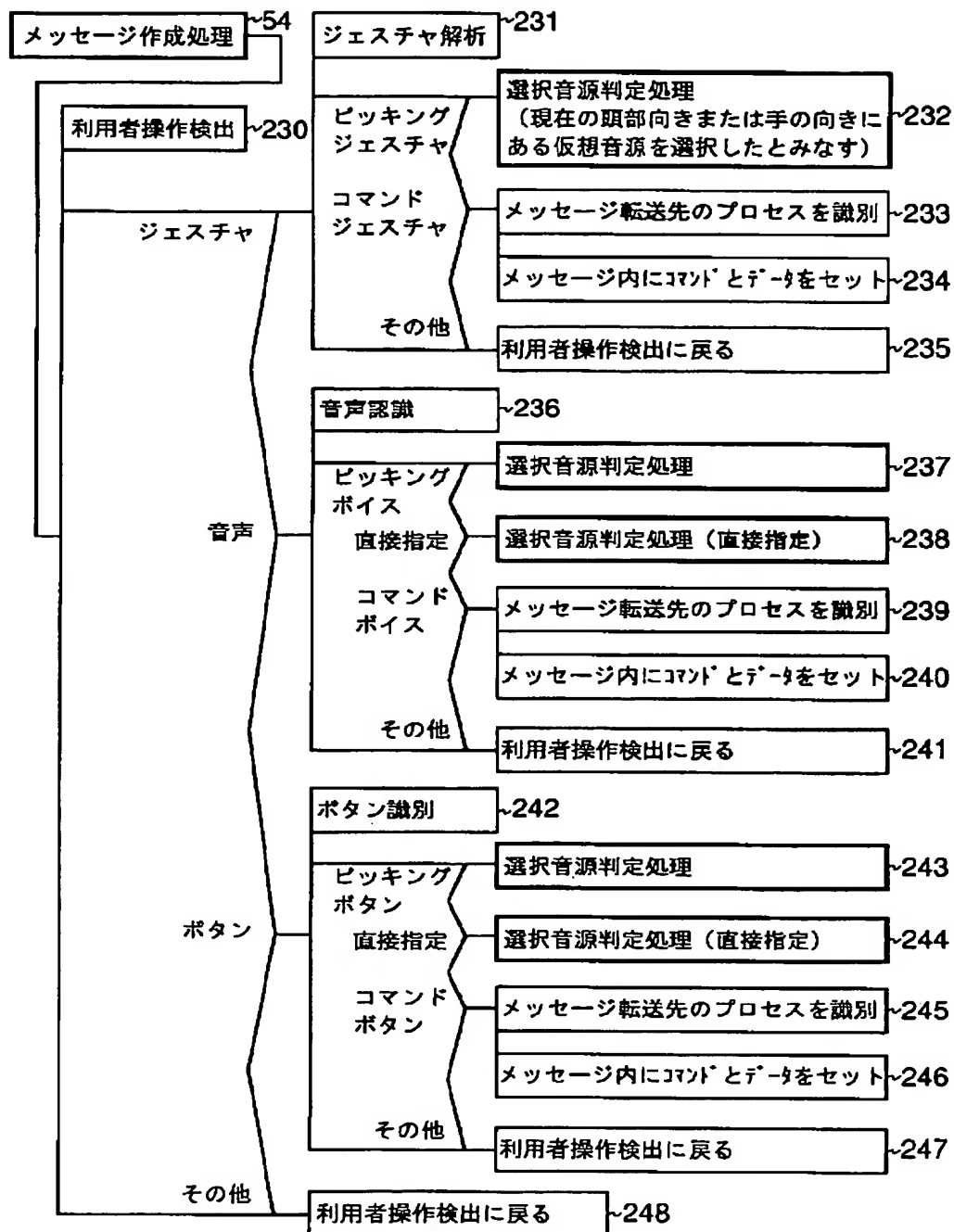
【図15】



(21)

【図11】

図11



(22)

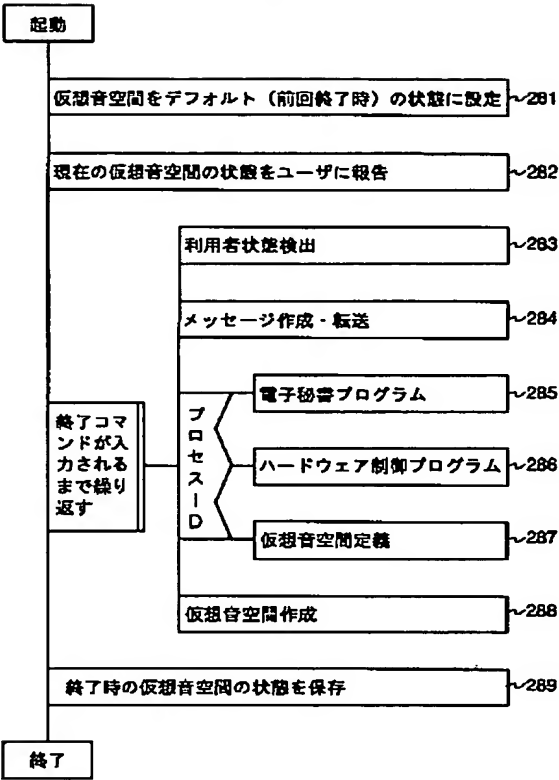
【図16】

図16

仮想音源位置	ヘッドホン出力	左	右
左		フィルタ1-1	フィルタ2-1
左前方		フィルタ1-2	フィルタ2-2
右前方		フィルタ1-3	フィルタ2-3
右		フィルタ1-4	フィルタ2-4

【図17】

図17



【図18】

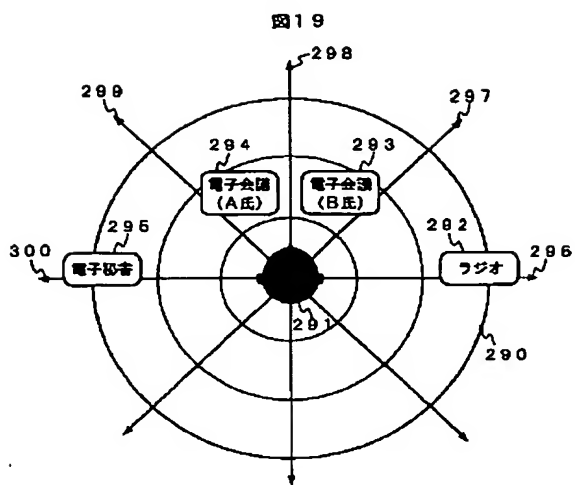
時刻		t1	t2	t3	t4	t5	t6	t7	
音情報選択		右	前方	左	前方	左	左前方	前方	右
入力（命令、データ等）			電子会議開始	電子秘書起動、命令入力	電子会議再開	電子秘書の処理結果入手	A氏と電子会議再開	B氏とも電子会議再開	電子会議終了
出力	左		電子秘書						
	左前方		電子会議（A氏）						
	右前方		電子会議（B氏）						
	右	ラジオ							

図18

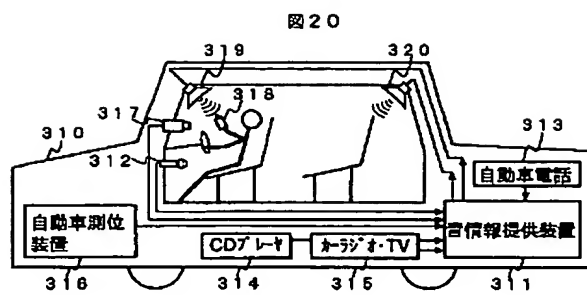
（※）図中の実線は音出力状態、点線は一時停止状態

(23)

【図19】



【図20】



【図21】

